

# Correlation-based model of artificially induced plasticity in motor cortex by a bidirectional brain-computer interface

Guillaume Lajoie,<sup>1</sup> Nedialko I. Krouchev,<sup>2</sup> John F. Kalaska,<sup>3</sup>  
Adrienne L. Fairhall,<sup>1,4,5</sup> and Eberhard E. Fetz<sup>1,4</sup>

<sup>1</sup>*University of Washington Institute for Neuroengineering, Seattle, Washington, USA*

<sup>2</sup>*Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada*

<sup>3</sup>*Groupe de recherche sur le système nerveux central,*

*Département de physiologie, Université de Montreal, Montreal, Quebec, Canada*

<sup>4</sup>*University of Washington, Dept. of Physiology and Biophysics, Seattle, Washington, USA*

<sup>5</sup>*University of Washington, Dept. of Physics, Seattle, Washington, USA*

Experiments show that spike-triggered stimulation performed with Bidirectional Brain-Computer-Interfaces (BBCI) can artificially strengthen connections between separate neural sites in motor cortex (MC). What are the neuronal mechanisms responsible for these changes and how does targeted stimulation by a BBCI shape population-level synaptic connectivity? The present work describes a recurrent neural network model with probabilistic spiking mechanisms and plastic synapses capable of capturing both neural and synaptic activity statistics relevant to BBCI conditioning protocols. When spikes from a neuron recorded at one MC site trigger stimuli at a second target site after a fixed delay, the connections between sites are strengthened for spike-stimulus delays consistent with experimentally derived spike time dependent plasticity (STDP) rules. However, the relationship between STDP mechanisms at the level of networks, and their modification with neural implants remains poorly understood. Using our model, we successfully reproduces key experimental results and use analytical derivations, along with novel experimental data. We then derive optimal operational regimes for BBCIs, and formulate predictions concerning the efficacy of spike-triggered stimulation in different regimes of cortical activity.

## INTRODUCTION

The cerebral cortex contains interacting neurons that form new functional connections through repeated activation patterns. For example, motor and somatosensory cortices are typically organized into somatotopic regions in which localized neural populations are associated with muscles or receptive fields and show varied levels of correlated activity (e.g. [9, 12, 21, 34, 42]). Functional relationships between such neural populations are known to change over time and reinforce relevant pathways [4, 29, 32, 41]. These changes are the result of plasticity mechanisms acting on myriad synaptic connections between cortical neurons. Most of them are relatively weak but can potentiate under the right conditions. However, it is not always clear what such conditions might be, or how one can interact with them for experimental or clinical purposes. Unanswered questions include the way local synaptic plasticity rules lead to stable, emergent functional connections, and the role of neural activity—and its statistics—in shaping such connections. While recent and ongoing work elucidates various plasticity mechanisms at the level of individual synapses, it is still unknown how these combine to shape the recurrently connected circuits that support population-level neural computations.

Bidirectional brain-computer interfaces (BBCI) capable of closed-loop recording and stimulation have enabled targeted conditioning experiments that probe these issues. In a seminal experiment [22], a BBCI recorded action potentials of a neuron at one MC site (labeled  $N_{\text{rec}}$ ) and delivered spike-triggered stimuli at another ( $N_{\text{stim}}$ ) for prolonged times in freely behaving macaque monkeys. This conditioning was able to increase functional connectivity from  $N_{\text{rec}}$  to  $N_{\text{stim}}$  (c.f. [39]), as measured by electromyogram (EMG) of muscle activation evoked by intracortical microstimulation (ICMS) in MC. Importantly, the relative strength of induced changes showed a strong dependence on the spike-stimulus delay, consistent with experimentally derived excitatory spike-time-dependent plasticity (STDP) time windows [2, 6, 28]. The effects of this protocol were apparent after about a day of ongoing conditioning, and lasted for several days afterwards. Similar spike-triggered stimulation showed corticospinal connections could increase or decrease, depending on whether the postsynaptic cells were stimulated after or before arrival of the presynaptic impulses [35]. This BBCI protocol has potential clinical uses for recovery after injuries and scientific utility for probing information processing in neural circuits.

The observations outlined above suggest that STDP is involved in shaping neural connections by normal activity during free behavior, and is the central mechanism behind the success of spike-triggered conditioning. However, this could not be verified directly as current experiments only measure functional relationships between cortical sites. Furthermore, interactions between BBCI signals and neural activity in recurrent networks are still poorly understood, and it remains unclear how BBCI protocols can be scaled up, made more efficient, and optimized for different experimental paradigms. For example, during spike-triggered stimulation, the spikes from a single unit are used to trigger stimulation of a neural population. While STDP can explain how the directly targeted synapses may be affected (i.e. from the origin of the recorded spikes to the stimulated population), the observed functional changes must rely on a broader scope of plastic changes involving other neurons that are functionally related to the recorded ones. What are the relevant network mechanisms that govern these population-level changes? How can a BBCI make use of population activity to trigger optimal stimulation?

Here we advance a modeling platform capable of capturing the effects of BBCI on recurrent, plastic neuronal networks. Our goal is to use the simplest dynamical assumptions in a “bottom-up” approach, motivated by neuronal and synaptic physiology, that enable the reproduction of key experimental findings from [22] at the functional level in MC and related muscles. In turn, we seek to use this model to provide insights into plasticity mechanisms in recurrent MC circuits (and other cortical regions) that are not readily accessible experimentally, as well as establish a theoretical framework upon which future BBCI protocols can be developed.

We build on a well-established body of work that enables analytical estimates of synaptic changes based on network statistics [5, 13–17, 24] and compare theoretical results with experiments and numerical simulations of a probabilistic spiking network. In our model, every neuron is excitatory — the modulatory role of inhibition in MC is instead represented implicitly by non-homogeneous probabilistic activation rates. While inhibition likely plays an important role in cortical dynamics, we consider results from our exclusive use of excitation to be a significant finding, suggesting that a few key mechanisms can account for a wide range of experimental results. Using data from previous work as well as from novel experiments, we calibrate STDP synaptic dynamics and activity correlation timescales to those typically found in MC neural populations. The result is a spiking model with multiplicative excitatory STDP and stable connectivity dynamics which can reproduce three key experimental findings:

- a. Using a single scaling parameter, we capture both (different) timescales of onset and dissipation of conditioning effects.
- b. With a simple filtered readout of our model’s output, we reproduce conditioning effects as measured by muscle EMG recordings evoked by ICMS.
- c. We reproduce the overall conditioning dependence on spike-triggered stimulation delay.

Furthermore, we make the following novel findings:

1. The distinct temporal statistics that characterize different regimes of cortical activity have an important impact on the efficacy of BBCI protocols.
2. Multi-synaptic mechanisms can lead to changes in circuit connectivity that are not predicted by STDP of synapses directly targeted by conditioning protocols.
3. When stimulating a subset of a given neural population, we find that the overall efficacy of conditioning depends supra-linearly on the proportion of this subset.

Together, these results outline quantifiable experimental predictions. They combine into a theoretical framework that is easily scalable and serves as a potential testbed for next-generation applications of BBIs. We discuss ways to make use of this framework in state-dependent conditioning protocols.

## RESULTS

Our model’s key assumptions are: (i) we only consider interactions between excitatory neurons; we revisit inhibition in the discussion; (ii) neurons in MC are sparsely and randomly

connected by synapses that are plastic and follow a single multiplicative STDP rule; (iii) the spiking activity of each neuron is modeled as a probabilistic process modulated by two net inputs: prescribed “external commands” —representing functional afferents to MC— and other neurons in the network via synaptic interactions.

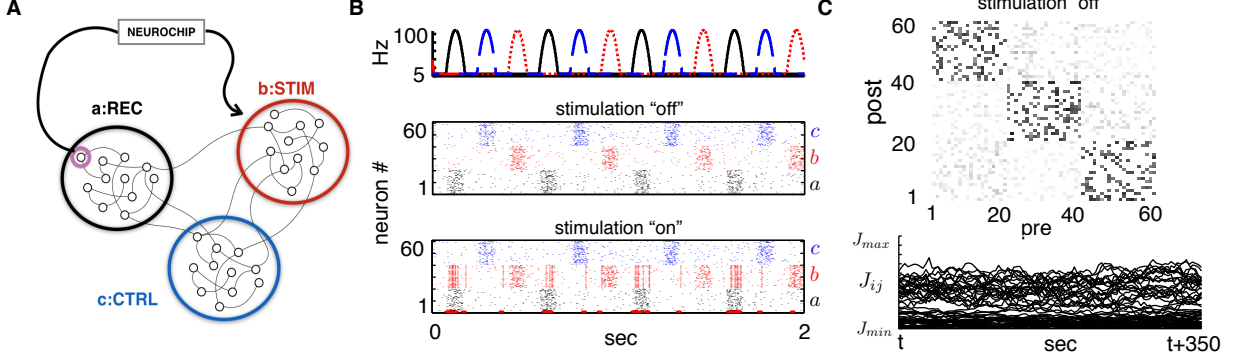


FIG. 1. (A) Diagram of network with three functional groups ( $a$ ,  $b$ ,  $c$ ). In the stimulation protocol ( $\dagger$ ), the “neurochip” records spikes from a single neuron in group  $a$  and delivers a large stimulus to group  $b$  after a delay  $d^\dagger$ . (B) Top: external rates  $\nu_a(t)$  (solid black),  $\nu_b(t)$  (dotted red) and  $\nu_c(t)$  (dashed blue) delivered to each neural group. Middle: sample spike raster plot from a simulation subject to rates plotted above with sparse, homogeneous connectivity  $J(0)$  ( $N = 60$ ). Bottom: same as middle but with spike-triggered stimulation ( $\dagger$ ) turned “on”. Spike from neuron number 1, circled in red, trigger the stimulation of all of population  $b$  (in red) after a delay of  $d^\dagger = 10$  ms. (C) Top: snapshot of connectivity matrix  $J(t)$  for network in B after spiking simulation of 5000 seconds (grayscale: white =  $J_{min}$ , black =  $J_{max}$ ). Bottom: example time traces of individual synapses from the same simulation.

We use a standard probability-based network model in which the instantaneous spiking probability rate  $\lambda_i(t)$  of neuron  $i$  changes in time and depends linearly on an external command  $\nu_i(t)$ , as well as the exponentially filtered spikes of other neurons in the network. Synaptic weights modulate the latter inputs and are stored in the connectivity matrix  $J(t)$  which is allowed to evolve in time according to a multiplicative STDP rule  $W(\Delta t, J_{ij})$  (see *Methods*). Here,  $\Delta t$  is the inter-spike-interval between a pre-synaptic spike from neuron  $j$  and a post synaptic spike from neuron  $i$ . Connectivity is chosen to be sparse and is initially drawn at random between the  $N$  neurons in the network with no new synapses allowed to emerge. Both axonal and dendritic spike transmission delays are sampled randomly for each synapse from two respective distributions with means  $\bar{d}^a$  and  $\bar{d}^d$ . For numerical simulations, each neuron is treated as an independent Poisson process with rate  $\lambda_i(t)$ . We refer the reader to *Methods* for more details about the model and its implementation.

To model the BBCI conditioning experiments, we divide our network of MC neurons into three groups of equal sizes as illustrated in Figure 1 A. Neurons in each group  $\alpha = a, b, c$  receive external inputs that have a group-specific modulating rate  $\nu_\alpha(t)$  but are otherwise independent from neuron to neuron (except for synaptic interactions). We drive the network with idealized input rates in all numerical simulations. These are composed of truncated sinusoidal functions offset with an equal phase difference between each group (Figure 1 B), and produce mean firing rates on the order of 10 Hz for most connectivity matrices. We later

introduce more biologically relevant statistics for activation rates to match simple model attributes to experimental data. To mimic the spike-triggered stimulation experiments in [22], we assign group  $a$  to be the “Recording” site, group  $b$  to be the “Stimulation” site and  $c$  to be the “Control” site. Neuron  $i = 1$  from group  $a$  is the “recorded” neuron, whose spikes trigger stimulation of every neuron in group  $b$  after a stimulation delay  $d^\dagger$ . Stimulation of subsets of  $b$  is also considered.

The relationship between spiking and synaptic activity in similar models of plastic networks has been extensively studied and predictions of synaptic dynamics under a variety of contexts were analytically derived (see e.g. [5, 13–17, 19, 24, 31, 33, 40, 48, 49, 51]). We build on these results and contribute novel findings that enable representations of artificial connections from BBCIs in generalized activity regimes. Our approach largely follows that of [17] but contains key innovations which are described in detail in *Methods*.

For illustration, Figure 1 C shows a snapshot of a connectivity matrix for a network of size  $N = 60$  driven by the rates shown in Figure 1 B and spike-triggered stimulation turned “off”, as well as traces of synaptic weights  $J_{ij}(t)$  evolving in time. It is evident that some structure emerges due to external commands, and that individual synapses show ongoing fluctuations surrounding that structure. Below, we often make use of averaged synaptic matrices  $\bar{J}(t)$ :  $3 \times 3$  matrices that contain group-averaged synaptic weights. In *Methods*, we derive analytical expressions for them, allowing us to forego lengthy spiking simulations. These expressions make use of group-averaged cross-correlation functions of different quantities that we often use throughout: for the external command rates ( $\hat{C}(u)$ ), for the network spiking activity without ( $C(u)$ ) and with spike-triggered stimulation ( $C^\dagger(u)$ ).

### Emergent synaptic structure and impact of spike-triggered stimulation

When neurons in the network’s three groups  $a$ ,  $b$ ,  $c$  are subject to external commands  $\nu(t) = (\nu_a(t), \nu_b(t), \nu_c(t))$  with stationary statistics, their averaged connectivity evolves toward an equilibrium that reflects these inputs’ correlations (c.f. [19, 31]), although individual synapses may continue to fluctuate. This was previously observed in a number of theoretical studies (see e.g. [17, 33]) and is consistent with the formation and dissociation of muscle assemblies in MC due to complex movements that are regularly performed [4]. The mean synaptic equilibrium  $\bar{J}^*$  strongly depends on the external inputs  $\nu(t)$ ’s correlation structure  $\hat{C}(u)$  (see Figure 2 A). Indeed, a narrow peak near the origin for correlations within groups, as is the case for the periodic external rates shown in Figure 1 B, along with the absence of such peaks for cross-group correlations, contribute to strengthening synapses within groups and weakening those across groups. Under such conditions, what will be the impact of spike-triggered stimulation?

Figure 2 B shows the evolution of synaptic averages  $\bar{J}_{\alpha\beta}(t)$ , analytically computed (see Figure 9 C in *Methods*) for a system initiated at the synaptic equilibrium associated with external rates  $\nu(t)$  from Figure 1 B. The inset of Figure 2 B shows the evolution of individual synapses from group  $a$  to group  $b$  from full network simulations. At 15 hours, the spike-triggered stimulation protocol is turned “on”, with a set delay  $d^\dagger = 20$  milliseconds, and synapses start changing. In  $\sim 10$  hours they reach a new equilibrium which differs from the initial one in a few striking ways, as seen in Figure 2 C and D, where normalized differences  $(\bar{J}^\dagger - \bar{J})/J_{max}$  are plotted for all pre- and post-group combinations. First, as expected and in accordance with experiments [22], the mean strength of synapses from group  $a$  ( $N_{rec}$ ) to

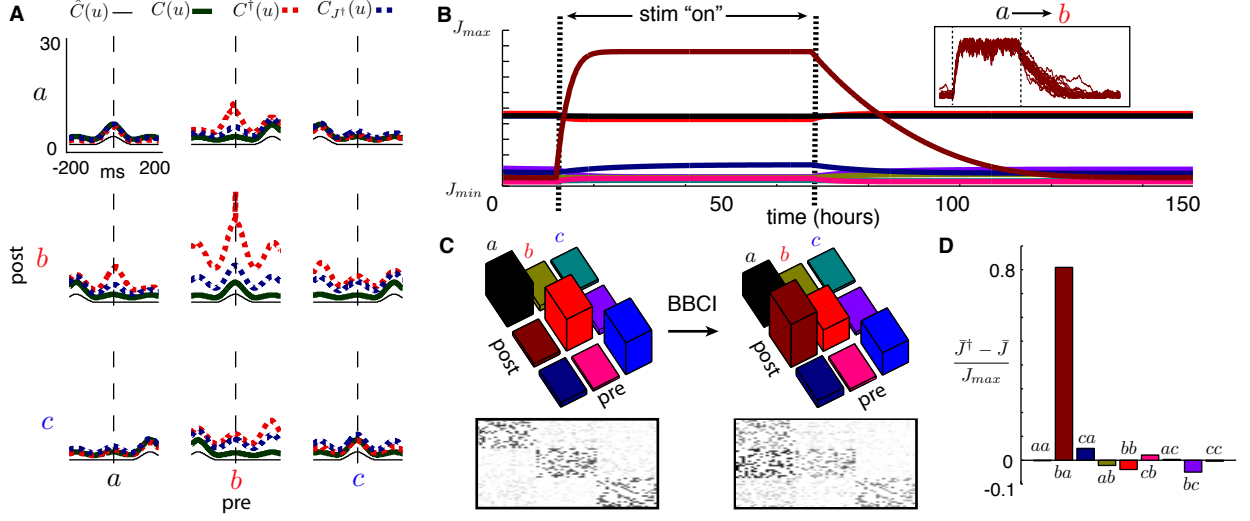


FIG. 2. (A) Mean network correlations for synaptic equilibria under normal activity  $C(u)$  (thick green line) and under spike-triggered stimulation  $C^\dagger(u)$  (dashed red line). Mean correlations for normal network activity with synaptic equilibrium obtained from stimulation ( $C_{J^\dagger}(u)$ ) is shown in dashed blue. External correlations  $\hat{C}(u)$  also shown (thin black line). (B) Evolution of synaptic weight averages  $\bar{J}_{\alpha\beta}(t)$  over time, computed with analytical estimates, color-coded as in B. External rates as in Figure 1 B. Initiated at equilibrium (see Figure 9 in *Methods*), spike triggered stimulation ( $\dagger$ ) is switched “on” for the indicated period. Inset shows evolutions of 10 randomly chosen synapses from group  $a$  to group  $b$ , from corresponding spiking network simulation. (C) Top: plots of equilibrium matrices  $\bar{J}^*$  in normal activity and under spike-triggered stimulation (left/right resp.). Bottom: snapshot of full simulated matrix  $J(t)$  once equilibria are reached. (D) Plot of relative synaptic changes between equilibria:  $(\bar{J}_{\alpha\beta}^\dagger - \bar{J}_{\alpha\beta})/J_{\max}$  with  $\alpha, \beta \in \{a, b, c\}$ .

group  $b$  ( $N_{stim}$ ) are considerably strengthened (by about 80%). As described in more detail below, this massive potentiation relies on two ingredients: correlated activity within group  $a$  and an appropriate choice of stimulation delay  $d^\dagger$ .

Perhaps more surprising are collateral changes in other synapses although they are of lesser magnitude. While this was previously unreported, it is consistent with unpublished data from the original spike-triggered stimulation experiment [22]. It is unclear how many of these changes are due to the particular external rate statistics and other model parameters; we return to this question below when realistic activity statistics are considered.

We also show that spike-triggered stimulation induces novel correlation structures due to synaptic changes as illustrated in Figure 2 A, which plots the correlations:  $\hat{C}(u)$ ,  $C(u)$ ,  $C^\dagger(u)$  and  $C_{J^\dagger}(u)$ . The latter denotes the correlations one observes under baseline activity (i.e. without ongoing stimulation) but with the equilibrium connectivity obtained after spike-triggered stimulation ( $\bar{J}^*$ ), i.e., at the end of spike-triggered stimulation. It is clear that every interaction involving group  $b$  is considerably changed, most strongly with group  $a$ , including the neuron used to trigger stimulation. More surprising is the increased cross-correlations of group  $b$  with itself, even though connectivity within group  $b$  is not explicitly potentiated by conditioning. In fact, it is slightly depressed (Figure 2 D). This happens because connections from group  $a$  to group  $b$  are considerably enhanced, which causes the mean firing rate of group  $b$  to grow and its correlations to increase. Later, we explore



similar collateral changes that occur because of multi-synaptic interactions. We see in the next section how these correlations translate into functional changes in the network.

Finally, Figure 2 B shows a crucial feature of our model: there are two timescales involved in the convergence from the normal to the artificial equilibrium, and the decay back to normal equilibrium after the end of spike-triggered stimulation. In the conditioning experiments from [22, 35], the effect of spike-triggered stimulation was seen after about 5-24 hours of conditioning, while the changes decayed after 1 to 7 days. With the simplified external drives producing a reasonable mean firing rate of about 10Hz for individual cells, an STDP learning rate of  $\eta = 10^{-8}$  was adequate to capture the two timescales of synaptic changes.

Thus, the simple excitatory STDP-based mechanisms in our model produce distinct timescales for increases and decay of the strength of synaptic connections produced by spike-triggered conditioning, in agreement with experimental observations. This does not contradict the findings from human psychophysical studies that feedback error-driven motor adaptation may involve two or more different and independent parallel processes with different temporal dynamics for learning and decay of the motor skill [45]. Nevertheless, the cellular mechanisms in our model may have some relation to the different timescales proposed to underlie motor adaptation at the system level [45]. Such relationships could be further investigated by direct experimentation and appropriate simulations.

In summary, our model satisfies the first experimental observation from [22] we set out to reproduce (point a. in *Introduction*). Indeed, we find that *two distinct timescales of synaptic changes (during and after conditioning) are an emergent property of our model, and tuning a single parameter is sufficient to reproduce the rates observed in experiments.*

### Simulating ICMS protocols: recovering functional changes

Changes in correlations due to spike-triggered stimulation indicate that there is an activity-relevant effect of induced synaptic changes, which is measurable from spiking statistics (see Figure 2 A). We now show how this is directly observable in evoked activity patterns that are consistent with ICMS protocols employed in experiments. In [22], connectivity changes were inferred using intra-cortical microstimulation (ICMS) and electromyogram (EMG) recordings of the monkey’s wrist muscles, as well as evoked isometric torques. To summarize, a train of ICMS stimuli lasting 50 ms was delivered to each MC site; simultaneously, EMG activity in three target muscles were recorded. The average EMG responses for repeated trials were documented for each of three MC sites (i.e. group *a*, *b* and *c*) before and after spike-triggered conditioning.

The experiment showed that prior to conditioning, ICMS stimulation of any MC site elicited well-resolved average EMG responses, largest in one muscle but not the two others. After conditioning, ICMS stimulation of the recording site ( $N_{rec}$ , group *a*) not only elicited an EMG response in its associated muscle, but also in that of stimulated site ( $N_{stim}$ , group *b*). While it was conjectured that synaptic changes in MC were responsible for the changes, this could not be verified directly. Our model suggests that synaptic changes can indeed occur in MC-like networks, but it remains unclear if such changes can lead to the experimentally observed motor output changes. We address this by simulating EMG responses of our model, before and after spike-triggered conditioning (†).

Figure 3 shows a simulated ICMS protocol before (panel A) and after (panel B) spike-triggered stimulation conditioning. For each case, synaptic matrices are chosen from full

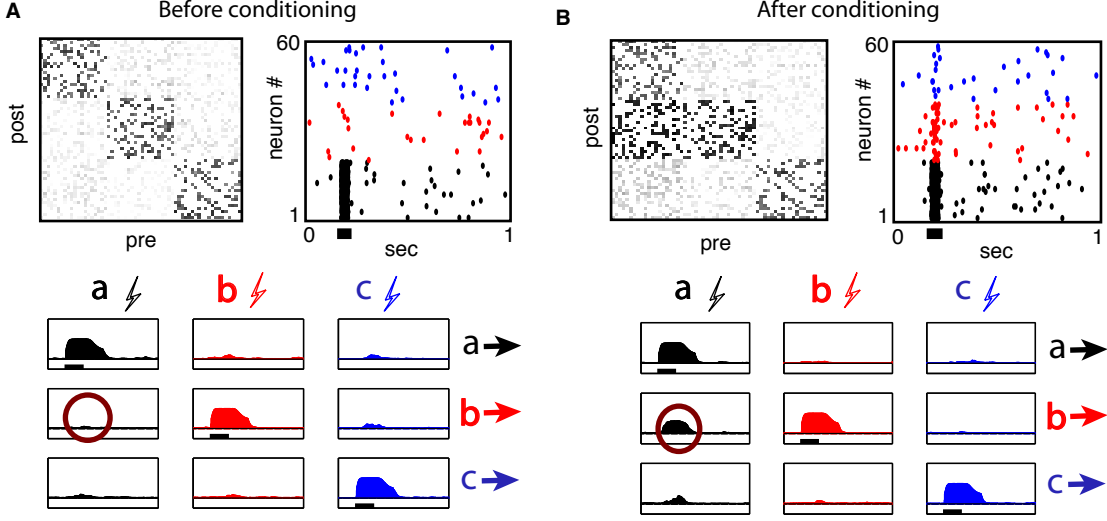


FIG. 3. (A) ICMS with baseline connections. (B) ICMS after spike-triggered stimulation. Top left panel: Connectivity matrix  $J$ . Top right panel: Raster plot of entire network as in Figure 1. Bottom black bar shows a 50 ms stimulation of all neurons in group  $a$  at 100 Hz. Bottom: Filtered responses of group projections. Columns designate neural groups being stimulated and rows designate the output evoked from each group that is equivalent to associated EMG responses.

network simulations and fixed (STDP is turned off), as shown in the top row of Figure 3 (reproduced from Figure 2 C). To mimic the ICMS stimulus, we add a square-pulse input rate of 100 Hz lasting 50 ms to the external rate  $\nu_\alpha(t)$  of a target group. An example of the spiking output of our network for  $\alpha = a$  is shown in the top row of Figure 3 where the solid black bar below the graph shows the stimulus duration. Next, we filter the spike output of all neurons within a group using the synaptic filter  $\epsilon(t)$  described earlier, and add them to obtain population activity time-courses. Finally, we take these summed profiles and pass them through a sigmoidal non-linearity  $-(1 + \exp -a(x - b))^{-1}$  with  $x$  the filtered activity—meant to represent the transform of neural activity to EMG signals. Here, we assume that the hypothetical motoneurons whose EMG is recorded receive inputs only from a single neural group and that network interactions are responsible for cross-group muscle activation.

We choose the nonlinearity parameters  $a = 2.5$  and  $b = 5$  to qualitatively reproduce the EMG responses seen in the experiment before spike-triggered conditioning: namely, well-resolved EMG responses are observed only when the relevant MC group is stimulated. The bottom row of Figure 3 shows simulated EMG responses of each group, averaged over 15 trials each, when ICMS stimulation is delivered to a single group at a time. Panel A shows little cross-group response to ICMS stimulation before spike-triggered stimulation conditioning. However, after conditioning, stimulation of group  $a$  evokes an emergent response in the muscle activated from group  $b$  (see circles in Figure 3), as well as a small increase for group  $c$ .

These features were both present in the original experiments (see Figure 2 in [22]) and are consistent with the synaptic strengths across groups before and after conditioning. In addition, the output from ICMS was also measured with isometric torques were produced in directions determined by the evoked EMG responses. The newly evoked EMG responses after



conditioning agree with the observation that torques evoked from  $N_{rec}$  (group  $a$ ) typically changed toward those previously evoked from  $N_{stim}$  (group  $b$ ). As such, from now on we equate an increase in mean synaptic strength  $\bar{J}_{\alpha\beta}$  between groups to an increase in functional connectivity.

We conclude that our model satisfies the second experimental observation from [22] we set out to reproduce (point b. in *Introduction*). That is, *a simple interpretation of evoked network activity —a filtered output of distinct neural group spiking activity— is consistent with the functional changes in muscle activation in ICMS protocols observed before and after conditioning.*

### Conditioning effects modulated by realistic activation statistics

Up to now, we used toy activation profiles  $\nu(t)$  in the form of truncated sinusoidal bumps to drive neural activity (Figure 1 B). In this section, we modify our simple model to incorporate experimentally observed cross correlation functions, whenever possible, in an effort to eliminate artificial activation commands and capture more realistic regimes. As a result, we no longer rely on numerical simulations of spiking activity, but rather on analytically derived averaged quantities to explore a wide range of conditioning regimes.

Before discussing spiking statistics, we note an important advantage of only considering mean synaptic strengths  $\bar{J}(t)$ . For spiking simulations, we use networks of  $N = 60$  neurons with probability of connection  $p = 0.3$ , which are considerably far from realistic numbers. Nevertheless, the important quantity for mean synaptic dynamics is the averaged summed strengths of synaptic inputs that a neuron receives from any given group:  $\frac{pN}{3}\bar{J}_{\alpha\beta}$ . Notice that many choices of  $p$  and  $N$  can lead to the same quantity, therefore creating a scaling equivalence. Moreover, additional scaling of  $J_{max}$  can further accommodate different network sizes. So far, we assumed that every neuron receives an average of 6 synapses from each group. If each of these synapse were at maximal value  $J_{max} = 0.1$ , then simultaneous spiking from a pre-synaptic group would increase the post-synaptic neuron’s spiking probability by 60%, a number we consider reasonable.

#### *Realistic activation statistics*

The changing dynamics of MC neurons in behaving macaques (e.g., [7, 10, 37, 50]) yield cross-correlations between cell pairs whose preferred feature can change depending on activity context [23, 46, 47]. We do not attempt to capture these subtleties with our simplified model. Rather, we create a family of cross-correlation functions  $\hat{C}(u)$  that roughly match averaged quantities observed experimentally, in an attempt to extract qualitative mechanisms. We assume, as was the case so far, that all mean cross-correlations are identical for neurons within the same group and neurons across groups respectively ( $\hat{C}_{\alpha\alpha}(u) = \hat{C}_{\beta\beta}(u)$  and  $\hat{C}_{\alpha\beta}(u) = \hat{C}_{\gamma\kappa}(u)$ ).

To calibrate our model, we use estimates of two quantities from MC neuron recordings: the mean cross-correlation between two neurons in the same group and the mean cross-correlation between neurons in different groups. Estimates were obtained from a monkey implanted with a Utah array in primary MC and performing a 2D target tracking task while seated in a chair. See *Methods* for details.

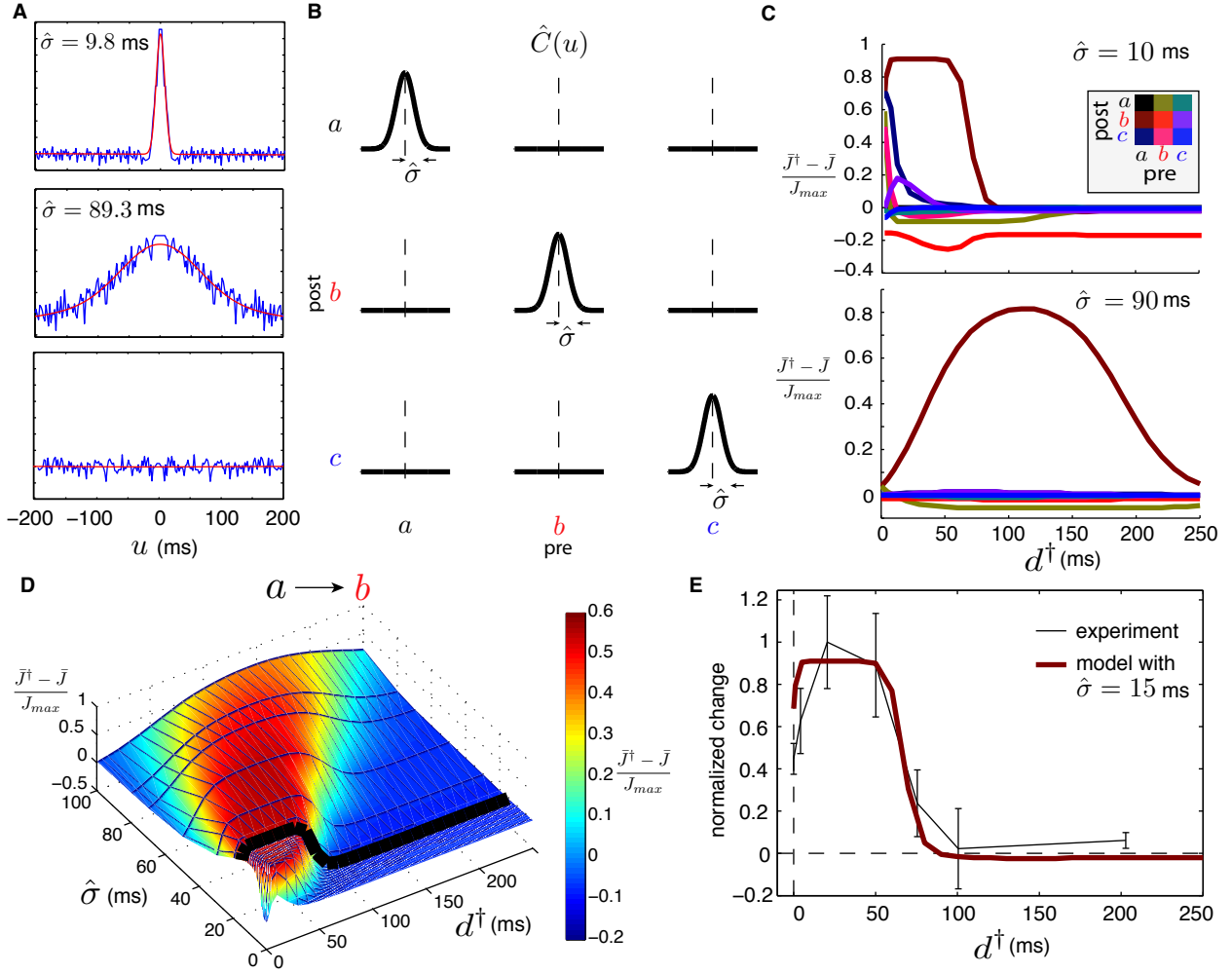


FIG. 4. (A) Experimentally obtained cross-correlograms of MC neurons in macaque monkey during a tracking task (blue) and Gaussian fit (red). Top, Middle: for neuron pairs recorded by the same electrode, respectively (using spike-sorting). Top shows the thinnest correlation peak ( $\hat{\sigma} = 9.8$  ms) and Middle shows the widest ( $\hat{\sigma} = 89.3$  ms). Bottom: For two neurons recorded by distinct electrodes. (B) Cartoon of Gaussian-shaped external cross-correlations  $\hat{C}(u)$ . (C) Relative differences  $(\bar{J}^\dagger - \bar{J})/J_{max}$  for all group-averaged synapses, as a function of stimulation delay  $d^\dagger$ , for the two extremal values of correlation width. Top:  $\hat{\sigma} = 9.8$  ms. Bottom:  $\hat{\sigma} = 89.3$  ms. (D) Relative differences for averaged synaptic strengths from group  $a$  to group  $b$  ( $(\bar{J}_{ba}^\dagger - \bar{J})/J_{max}$ ) as a function of stimulation delay  $d^\dagger$  and correlation peak width  $\hat{\sigma}$ . Black line corresponds to best fit plotted in E. (E) Superposition of relative difference for  $\bar{J}_{ba}$  and normalized mean torque change from spike-triggered conditioning experiments on macaque monkeys. Experimental data from Figure 4 of [22]; error bars show the standard error of the mean. Best fit between model and experimental curve is for  $\hat{\sigma} = 15$  ms (see black line in D).

In this data set, there were only a few pairs of neurons recorded from the same electrode. Nevertheless, such well-resolved pairs typically showed Gaussian-shaped cross-correlograms roughly centered at the origin. An important observation is that the width of the Gaussian peaks varied greatly. Figure 4 A top and middle show both extremes where the cross-correlograms were fitted to Gaussian functions with a standard deviation  $\hat{\sigma}$  ranging from

$\sim 10$  to  $\sim 90$  milliseconds. In contrast, neuron pairs from different electrodes showed more subtle structure in their cross-correlograms, and a wide variety of feature shapes (e.g. asymmetrical relationships, anti-correlations, etc.). Most of them showed no correlation features, as depicted by the typical flat cross-correlogram in the bottom of Figure 4 A. An important point in relating these observations to our model is that the mean cross-correlation, across many pairs of neurons within specific groups, is the defining factor that influences synaptic dynamics. While the examples shown in Figure 4 A are insufficient to estimate such averages, they are able to guide parameter ranges in our model.

In light of these observations, we make the following simplifications concerning the model's idealized cross-correlations  $\hat{C}(u)$ : cross-correlations of neurons within a group are Gaussian while those for neurons across groups remain flat. We get:

$$\hat{C}_{\alpha\beta}(u) = \begin{cases} \frac{L}{\hat{\sigma}\sqrt{2\pi}} \exp(-\frac{u^2}{2\hat{\sigma}^2}) + B & ; \alpha = \beta \\ B & ; \alpha \neq \beta \end{cases} \quad (1)$$

where  $B$  is a baseline,  $L$  is the height of the peak and  $\hat{\sigma}$  modulates the peak's width (i.e. standard deviation). Figure 4 B illustrates our idealized cross-correlation function  $\hat{C}(u)$ . From the analysis in *Methods*, it can be shown that the parameters  $B$  and  $L$ , along with  $\eta$ , only influence the speed at which synaptic changes occur (see Equation (24)). Therefore, we set these to obtain a mean firing of about 10 Hz as for the previously assumed activation rates  $\nu(t)$ , along with  $\eta$  to obtain convergence timescales similar to those shown in Figure 2 B. We henceforth set  $B = 0.17$ ,  $L = 2.13$  and  $\eta = 10^{-8}$ . Since experiments show that  $\hat{\sigma}$  can take a range of values, we leave it as a free parameter in our regime exploration below.

It should be noted that experimentally measured cross-correlations represent the activity of the recorded neurons themselves, as opposed to the cross-correlations of external commands ( $\hat{C}(u)$ ) as we assume here. Nevertheless, we make the simplifying assumption that the resulting timescales of self-consistent network cross-correlations are similar enough to those of external drives to justify the form of  $\hat{C}(u)$  used here. Further comments on this assumption can be found in the *Discussion*.

#### *Cross-correlation width influences optimal stimulation delay*

We compute the relative change in mean synaptic strengths  $((\bar{J}^\dagger - \bar{J})/J_{max})$  before and after spike-triggered stimulation, for a range of cross correlation widths  $\hat{\sigma}$  and stimulation delays  $d^\dagger$ . Figure 4 C shows these differences, for all pre- and post-synaptic neural group combinations, for  $\hat{\sigma} = 10$  ms and  $\hat{\sigma} = 90$  ms. Notice that the correlation width has significant effects on the  $d^\dagger$ -dependence of BBCI conditioning for all synapses. For now, we concentrate on the synapses directly targeted by the stimulation procedure, namely those from group  $a$  to group  $b$  ( $\bar{J}_{ba}$ ), which naturally show the greatest changes. We discuss the changes incurred by other synapses below.

For both values of  $\hat{\sigma}$  used in Figure 4 C, the changes incurred by  $\bar{J}_{ba}$  depend non-monotonically on the triggering delay  $d^\dagger$ , admitting either a maximum or a plateau. However the maximizing  $d^\dagger$  value (or range) changes with  $\hat{\sigma}$ , and so does the sensitivity to choice of  $d^\dagger$ , i.e., how rapidly the curve reaches its maximum as  $d^\dagger$  is varied. Figure 4 D comprehensively illustrates this dependence, showing  $(\bar{J}_{ba}^\dagger - \bar{J}_{ba})/J_{max}$  for a range of both  $\hat{\sigma}$  and  $d^\dagger$ . The main features of this relationship are: (i) optimal delays ( $d^\dagger$  at peak) get larger as  $\hat{\sigma}$  grows. (ii)

sensitivity is mitigated by larger  $\hat{\sigma}$ , i.e., large correlation timescales lead to optimal delays whose conditioning outcomes are less sensitive to perturbations.

The relative simplicity of our model’s equations enables a mechanistic interpretation of these findings. Indeed, Equation (23) illustrates that it is the integral of the network’s correlations multiplied by the STDP rule that dictates synaptic dynamics. When artificial stimulation is delivered, components of this integral are shifted by  $d^\dagger$ . Wider correlations imply that small changes in shifts  $d^\dagger$  have slower effects on the integral, leading to both point (i) and (ii) listed above.

When the stimulation of group  $b$  is triggered on spikes from the recorded neuron in group  $a$ , the plasticity incurred by synapses from other neurons in group  $a$  to any neuron in group  $b$  will depend on how likely it was that  $a$ -neurons fired within a short interval of the recorded neuron’s spikes. Correlation width measures the synchrony of spiking activity. For narrow correlations, neurons in group  $a$  fire closely together and the delay required to induce the maximal plastic change directly depends on the combination of axonal delay and the off-width of the potentiation peak of the STDP rule (see Figure 7 A in *Methods*). Essentially, the peak of the cross-correlation needs to be shifted so it aligns with the peak STDP potentiation. For wider correlation functions, synaptic changes depend less on synchrony and more on interactions with larger time lags. Indeed, optimal delays are those that shift the cross-correlation’s peak past the STDP peak, so that the latter is aligned with the right-hand-side tails of cross-correlations. This relationship leads to more robust potentiation at larger stimulation delays.

This phenomenon constitutes our first novel finding (point 1. from *Introduction*): *neuronal activity statistics have an important impact on the value of optimal spike-triggered stimulation delays ( $d^\dagger$ ), and can lead to synaptic potentiation that is more robust deviations from that optimal value.*

### Qualitative reproduction of stimulation/plasticity dependence

It remains unclear if the mechanisms described above are consistent with the experimentally observed relationship between stimulation delay ( $d^\dagger$ ) and efficacy of spike-triggered conditioning in macaque MC. We investigate this by comparing efficacy, as measured by the percentage of torque direction change evoked by ICMS before and after conditioning [22], to relative synaptic strength changes in our model. This is motivated by the above demonstration that synaptic strengths are well correlated with amplitude of evoked muscle activations in a ICMS experiment (see Figure 3 and point b. in *Introduction*). Nevertheless, the following comparison between model and experiment is qualitative, and meant to establish a correspondence of (delay) timescales only.

We use the data originally presented in Figure 4 of [22], describing the shift in mean change in evoked wrist torque direction by ICMS of the  $N_{rec}$  site (analogous to our group  $a$ ), as a function of stimulation delay  $d^\dagger$ . We plot the same data in Figure 4 E, with the maximal change (in degrees) normalized to one. On the same graph, we plot the  $(\bar{J}_{ba}^\dagger - \bar{J}_{ba})/J_{max}$  v.s.  $d^\dagger$  curve for the value of  $\hat{\sigma}$  that offers the best fit (in L1-norm). This amounts to finding the best “ $\hat{\sigma}$ -slice” of the graph in Figure 4 D to fit the experimental data. We found that  $\hat{\sigma} \simeq 15$  ms gives the best correspondence.

We reiterate that this comparison is qualitative. Nevertheless, the fit between the  $d^\dagger$ -dependence of experimentally observed functional changes and modelled synaptic ones is

clear. As our model’s spiking activity and STDP rule are calibrated with experimentally observed parameters (see *Methods*), this evidence suggests that our simplified framework is consistent with experiments. Importantly,  $\hat{\sigma} = 15$  ms is comparable to correlation timescales between functionally related MC neurons in macaque during a stereotyped motor task [21, 46, 47], where such correlations are reported to be gaussian-like and have peak width at half height on the order of 20 ms, corresponding roughly to  $\hat{\sigma} = 10$  ms. It is generally believed that task-specific motion introduces sharp correlations and that free behaving, rest and sleep states induce longer-range statistics [23]. The conditioning experiment in [22] was conducted over a wide range of states, including sleep, which may lead to longer mean cross-correlation timescales. A prediction of our model is that spike-triggered conditioning restricted to periods of long timescale correlations in MC, such as during sleep [23], could lead to a more robust conditioning dependence on stimulation delays (see *Discussion*).

This finding implies that our model successfully reproduces the third and last experimental observation from [22] (point c. from *Introduction*): *using simplified cross-correlation functions of MC neural populations calibrated from experimental measurements, our model reproduces the relationship between the magnitude of plastic changes and the stimulation delay in a spike-triggered conditioning protocol.*

### Multi-synaptic and collateral effects.

We now explore the effects of spike-triggered stimulation on collateral synaptic strengths, i.e., other than the targeted  $a$ -to- $b$  pathway ( $N_{rec}$ -to- $N_{stim}$ ). Figure 5 shows color plots of these changes as a function of  $d^\dagger$  and  $\hat{\sigma}$ , for the nine combinations of pre- and post-synaptic groups.

For a wide range of parameters, there is little change other than for the  $a$ -to- $b$  synapses. Indeed, when external cross-correlation width  $\hat{\sigma}$  is moderate to large, spike-triggered stimulation has little effect on collateral connections, for any stimulation delay. This is in conjunction with the robustness of  $a$ -to- $b$  changes discussed in the previous section. Nevertheless, some localized features arise when the external cross-correlation width  $\hat{\sigma}$  is small.

First,  $b$ -to- $b$  synapses become depressed, regardless of stimulation delay, for short correlation timescales. This is due to the occurrence of synchronized population spiking produced by artificial stimulation which, because of our choice of dendritic and axonal delays (see *Methods*), promote depression. Such synchronized spikes induce sharp  $\delta$ -peaks in the network’s cross-correlation (see Figure 8 B in *Methods*) and the combination of transmission delays shifts this peak toward the depression side of the STDP rule. When external cross-correlations are narrow (i.e. small  $\hat{\sigma}$ ), their interaction with the STDP rule—which manifests in the integral in Equation (13)—is more sensitive to the addition of such  $\delta$ -peaks, resulting in overall depression. In contrast, when cross-correlations are wider, the addition of  $\delta$ -peaks has a smaller effect since a wider range of correlations contribute to the integrated STDP changes.

Second, the synapses from  $b$  to  $a$ , become potentiated for short delays  $d^\dagger$  when  $\hat{\sigma}$  is small enough. This happens because of a combination of factors. When the recorded neuron in group  $a$  spikes, the population-wide spike artificially elicited in  $b$  quickly follows and travels to the  $b$ -to- $a$  synapses. This means that the spike of a single neuron in  $a$  affects all neurons in  $a$ , with an effect amplified by the strength of many synapses, shortly after the neuron originally fired. When cross-correlations among  $a$ -neurons are wide, the effect of this

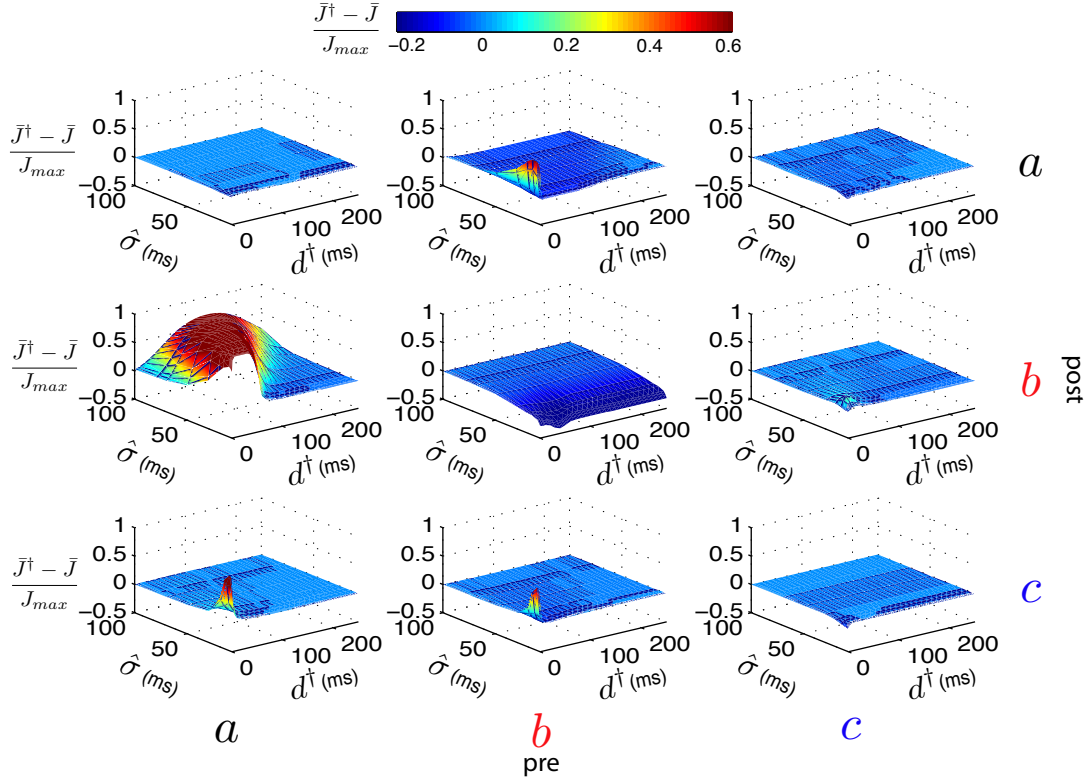


FIG. 5. Relative differences for mean synaptic strengths  $(\bar{J}^\dagger_{\alpha\beta} - J)/J_{max}$  as a function of stimulation delay  $d^\dagger$  and correlation peak width  $\hat{\sigma}$ , for all combinations of pre- and post-synaptic groups  $a$ ,  $b$ ,  $c$ .

mechanism is diluted, similarly to the  $b$ -to- $b$  synapses discussed above. However, when  $a$  is nearly synchronous, this short-latency feedback produces synaptic potentiation of the  $b$ -to- $a$  synapses.

Lastly, the synapses from both groups  $a$  and  $b$  onto the control group  $c$  are also potentiated when both  $\hat{\sigma}$  and  $d^\dagger$  are small enough. This can be explained in two parts and involves di- and tri-synaptic mechanisms. When the recorded neuron in  $a$  fires a spike, a population-wide spike is artificially evoked in  $b$  shortly after, which travels down to  $b$ -to- $c$  synapses and elicits a response from neurons in  $c$ . Narrow cross-correlations imply that many spikes in  $a$  fall within a favorable potentiation window of spikes in  $c$ , thereby contributing to the potentiation of  $a$ -to- $c$  synapses. In turn, the strengthening of  $a$ -to- $c$  synapses imply that spikes in  $a$  are more likely to directly elicit spikes in  $c$ , thereby repeating the same process in a different order for  $b$ -to- $c$  synapses. Note that without  $a$ -to- $c$  synapses, the  $b$ -to- $c$  synapses would not potentiate. Indeed, as was the case for  $b$ -to- $b$  synapses, the combination of transmission delays do not conspire to promote direct potentiation following a population-wide synchronous spike.

This illustrates that emergent, multi-synaptic mechanisms give rise to significant changes in collateral synapses, and that these mostly occur for short stimulation delays and short correlation timescales.

This constitutes our second novel finding (point 2. from the *Introduction*): *multi-synaptic mechanisms give rise to significant changes in collateral synapses during conditioning with short delays, and these cannot be attributed to STDP mechanisms directly targeted by the*



*BBCI, instead emerging from network activity.*

### Stimulating a subgroup of neurons

We have assumed that spike-triggered stimulation elicits population-wide synchronous spiking of all neurons in group  $b$  ( $N_{\text{stim}}$  in [22]). This is valid if the neural group  $b$  represents all the neurons activated by the stimulating electrode, but is not necessarily representative of the larger population of neurons that share external activation statistics due to a common input  $\nu_b(t)$ . Indeed, some neurons that activate in conjunction with those close to the electrode may be far enough from it so they do not necessarily spike in direct response to a stimulating pulse. Alternatively, selective activation of neurons within a group can also be achieved via optogenetic stimulation in a much more targeted fashion [25]. We now consider the situation in which only a certain proportion of neurons from group  $b$  is activated by the spike-triggered stimulus.

We denote the stimulated subgroup by  $b^\dagger$  and the un-stimulated subgroup by  $b^\circ$  (Figure 6 A). All neurons in group  $b$  receive the same external rates  $\nu_b(t)$  as before, but only a few (solid red dots) are stimulated by the Neurochip. Let  $N_b = N/3$  be the number of neurons in group  $b$  and the parameter  $\rho$ , with  $0 \leq \rho \leq 1$ , represent the proportion of stimulated neurons in  $b$ . The sizes of groups  $b^\dagger$  and  $b^\circ$  are given by  $N_b^\dagger = \rho N_b$  and  $N_b^\circ = (1 - \rho)N_b$ , respectively. We now adapt our analytical averaged model (23) to explore the effect of stimulation on subdivided synaptic equilibria. We verified that the analytical derivations used below match the full spiking network simulations as before (data not shown).

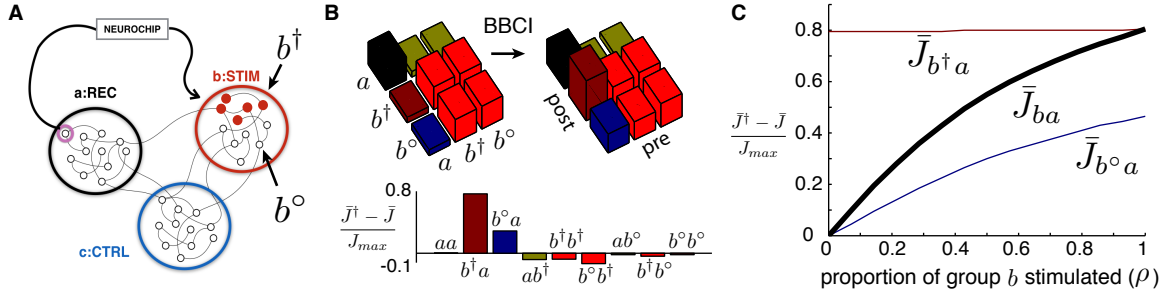


FIG. 6. (A) Network with group  $b$  subdivided into two subgroups: stimulated neurons ( $b^\dagger$ , red) and unstimulated neurons ( $b^\circ$ , white). (B) Top: plots of equilibrium synaptic strengths  $\bar{J}^*$  between group  $a$  and both subgroups of  $b$  in normal activity and under spike-triggered stimulation (left/right resp.). Proportion of stimulated neurons is set to  $\rho = 0.5$ . Bottom: Plot of relative synaptic changes between equilibria:  $(\bar{J}_{\alpha\beta}^\dagger - \bar{J}_{\alpha\beta})/J_{max}$ . (C) Relative synaptic changes as a function of proportion of neurons in group  $b$  receiving stimulation  $\rho$ . Plotted are subgroup averaged synaptic changes for synapses from  $a$  to  $b^\dagger$  and  $b^\circ$  and from  $a$  to entire group  $b$ .

Both subgroups of  $b$  receive the external rate  $\nu_b(t)$  but only one receives spike-triggered stimulation. These changes are captured in the averaged analytical derivations by tracking the number of neurons in each sub-group in the averaging steps leading to Equations (22) and (23) accordingly — replacing  $N/3$  by  $N_b^\dagger$  and  $N_b^\circ$  where necessary. This way, we obtain subgroup-specific synaptic averages (e.g.  $\bar{J}_{b^\dagger a}$ ).

Figure 6 B shows the group-averaged connectivity strength between group  $a$  (the recording site) and both subgroups of  $b$ , before and after spike-triggered stimulation. External

cross-correlations  $\hat{C}(u)$  are as in Figure 4 B with  $\hat{\sigma} = 20\text{ms}$ , and stimulation delay is set at  $d^\dagger = 30\text{ms}$ . The proportion of stimulated neurons in  $b$  is set to  $\rho = 0.5$ . The bottom of the same panel shows the normalized changes of mean synaptic strengths due to spike-triggered stimulation. As established for the original network (see Figure 2 B), the biggest change occurs for synapses from  $a$  to the subgroup that is directly stimulated ( $b^\dagger$ ). However for subgroup  $b^\circ$ , we see a noticeable change in its incoming synapses from group  $a$ , in contrast to synapses of other un-stimulated groups ( $c$ ) that do not appreciably change (not shown). This means that sharing activation statistics with stimulated neurons is enough to transfer the plasticity-inducing effect of conditioning to a secondary neural population.

Next, we investigate how this phenomenon is affected by the proportion of neurons in  $b$  that receive stimulation:  $\rho$ . Figure 6 C shows the subgroup-averaged normalized changes of synapses from group  $a$  to subgroups  $b^\dagger$  and  $b^\circ$ , as  $\rho$  varies between 0 and 1. When more neurons get stimulated, the transferred effect on the unstimulated group is amplified. This means that the combined outcome on the entirety of group  $b$  grows even faster with  $\rho$  (supra-linearly), as shown in Figure 6 C, where the combined  $b$ -averaged changes in synaptic strength,  $\bar{J}_{ba} = \rho\bar{J}_{b^\dagger a} + (1 - \rho)\bar{J}_{b^\circ a}$ , are plotted as a function of  $\rho$ .

In summary, this phenomenon represents our third and final finding (point 3. in *Introduction*). Our model shows that *neurons not directly stimulated during spike-triggered conditioning can be entrained into artificially induced plasticity changes by a subgroup of stimulated cells, and that the combined population-averaged effect grows supra-linearly with the size of the stimulated subgroup.*

## DISCUSSION

### Summary

In this study, we used a probabilistic model of spiking neurons with plastic synapses obeying a simple STDP rule to investigate the effect of a BBCI on the connectivity of recurrent cortical-like networks. Here the BBCI records from a single neuron within a population and delivers spike-triggered stimuli to a different population after a set delay. We developed a reduced dynamical system for the average synaptic strengths between neural populations; these dynamics admit stable fixed points corresponding to synaptic equilibria that depend solely on the network activity's correlations and spike-triggered stimulation parameters. In this framework, individual synapses may fluctuate with ongoing network activity but their population average remains stable in time. We validate our findings with detailed numerical simulations of a spiking network and calibrate our result based on experiments in macaque MC. To our knowledge, this is the first time a plastic spiking network model includes recurrent network interactions to capture the effects of a BBCI.

We successfully reproduce key experimental results from [22] that describe synaptic changes due to spike-triggered conditioning. Specifically, we recover the two emergent timescales with which these changes occur (hours to days), we show that filtered evoked activity from our network mimics muscle EMG evoked by ICMS protocols, and we show that maximal changes in mean synaptic strength from the recording site to the stimulated site occur with stimulation delays in the 20-50 ms range, as was the case for experiments. Furthermore, we formulate three novel findings using our model. We outline the relationship between temporal statistics within neural populations and optimal stimulation parameters,

we uncover multi-synaptic mechanisms that emerge from spike-triggered conditioning that are not directly predicted by a single-synapse STDP rationale, and we found that the stimulation of a subset of neurons within a population can lead to supra-linear scaling of effects.

Based on this, we formulate two main experimental predictions:

1. Spiking statistics in a network greatly influence the effects of stimulation, with a wider range of spike-stimulation delays leading to optimal evoked plasticity when correlations have timescales. We predict that restricting spike-triggered stimulation to states with these statistics, such as sleep, or potentially driving such correlation artificially, may provide more robust outcomes.
2. For very short delays, stimulation may produce collateral effects in other synaptic connections due to polysynaptic interactions. We propose that under special conditions, it may be possible to influence collateral synaptic pathways that are not directly targeted by a spike-triggered stimulation protocol.

### **Cortical activity: driven vs emergent**

In our model, input rates are needed to endow neural groups with desired statistics. In cortical networks it is unclear how much of observed activity is driven from external sources and how much is due to intra-network interactions. We argue that the use of prescribed external activations is appropriate since we show that plasticity works to “learn” activation correlations (see also [13–17]), thereby guiding synaptic connections to promote the same spontaneous network statistics as its driven ones. Therefore, while external input rates are necessary in our model, the resulting network correlations  $C(u)$  reflect both emergent and external dynamics, and serve as a proxy for MC activity.

### **State-dependent stimulation: designing adaptive stimulation strategies**

An interesting outcome of our findings is that spike-triggered conditioning is strongly influenced by the source and target cross- and auto-correlation structure  $C(u)$ . Throughout the paper, we assumed that these statistics were stationary for the entire simulated period. However, it is well known that cortical circuits can show a wide range of activity regimes depending on the state of the animal. The statistics of MC neurons may be very different if the animal is awake and behaving freely or performing a precise motor task, or is asleep [23]. While such states have limited durations, their timescales (from minutes to hours) may be long enough to define locally-stationary statistics that BCI protocols could leverage to optimize desired effects. For example, stimulation during sleep, which is known to produce oscillation-rich activity with longer-range correlations [1, 23], could have more robust but slower effects, while stimulation during a specific task can have more targeted outcomes.

### **Scalable framework: from experimental to clinical applications**

Our model can easily be scaled up to include multiple recording and stimulation sites, and different stimulation protocols, including, e.g. EMG-triggering [27] or paired-pulse stimulation [43]. It is also easily adaptable to optically-based stimulation which can target specific

neurons within functional groups (see e.g. [8, 25, 44]). As it does not require costly simulations —only theoretical estimates— it is straightforward to apply optimization algorithms to find the best stimulation protocol to achieve a desired connectivity between cortical sites. Furthermore, it can easily incorporate closed-loop signals such as changes in recorded statistics in real-time. This framework offers a flexible testbed to help design experiments and clinical treatments.

Neural implants such as BCIs and BBCIs are under active development as they have significant potential for clinical use [38]. Among other outcomes, they are promising avenues for treatment of motor disabilities. Indeed, a BBCI capable of inducing plastic changes in cortical circuits could be used to promote novel synaptic pathways in order to restore functional connectivity after a stroke or injury [11, 18].

For such bidirectional neural implants to be successful, a number of real-time computational issues need to be resolved. Our modeling framework presents an important step toward development of rapidly computable guides for controllers, based on anatomical organization, measurable network properties and known physiological mechanisms such as STDP.

### Model limitations and future steps

While being able to capture important features, our model misses some important physiological aspects of MC circuits. Nevertheless, we note that the simplicity of our model is a strength: it captures complex network-level plasticity changes with only excitatory activity and STDP. This suggests that excitatory mechanisms are central to artificially-induced plasticity by a BBCI. However, this simplicity will likely be insufficient to reproduce more complex protocols that also recruit inhibitory populations. An example is paired-pulse conditioning, where the BBCI stimulates several neural sites, with different time delays. Additional biological realism is key to expand our bottom-up theoretical framework.

A number of steps can be taken to add physiological realism, each of them adding some complications. For inhibition, there are technical issues when considering probabilistic spiking and inhibition (e.g. inhibition can induce “negative” spiking probabilities if unchecked), and STDP mechanisms for inhibitory synapses are not well understood. The implementation of our framework in a dynamical model setting, building on theoretical models of inhibitory plasticity (see e.g. [52]) are natural next steps. In addition, the inclusion of spatial structure and heterogeneous delay distributions, based on anatomical data [29], is necessary to expand the framework to multiple cortical sites and spinal cord. Finally, the inclusion of modulatory mechanisms, activity- or chemically-dependent, is crucial to capture phenomena such as synaptic consolidation and adaptation.

In summary, our reduced dynamical system approach is a promising basis upon which to build and ultimately to predict the effects of finer-grained cell-type specific and temporally structured activation patterns afforded by next-generation neural implants using electrical or optical stimulation.

## METHODS

This section contains details about our model, numerical simulations and analytical derivations of averaged synaptic dynamics. We closely follow prior literature [17], and de-

velop key innovations to better suit our model’s BBCI components.

## Model

### *Spiking activity*

We consider a network of  $N$  model neurons whose underlying connections are given by an  $N \times N$  connectivity matrix  $J(t)$  with synaptic weights that evolve according to a plasticity rule (as described below). The presence of a directed synapse  $J_{ij}$  from neuron  $j$  to neuron  $i$  is randomly and independently determined with probability  $p$ . Throughout this paper, we consider sparse networks with  $p = 0.3$ . Once an initial connectivity matrix  $J(0)$  is drawn, existing synapses are allowed to change but no new synapses can be created. All synapses are excitatory so that  $J_{ij} \geq 0$ .

The spike train of neuron  $i$  is a collection of times  $\{t_i^s\}$  and can be written as sum of Dirac  $\delta$  functions:  $S_i(t) = \sum_{t_i^s} \delta(t - t_i^s)$ . We are interested in the probability of observing a certain spike train  $S_i(t)$  given a certain connectivity matrix  $J(t)$ , the other neurons’ activity and the external driving signal  $\nu_i(t)$ . We assume that this spiking probability can be expressed as a time-dependent density  $\lambda_i(t)$ , the parameters of which are governed by the *ensemble average*  $\langle S_i(t) \rangle$  and must be consistent across the network. For simulations, we assume  $\lambda_i(t)$  is a Poisson rate, but most of the derivations we make below are generalizable to other probabilistic or dynamical models of network spiking activity. For example, an *integrate-and-fire* spiking mechanism could be used, with a stochastic transfer function derived from Fokker-Planck equations using linear response theory, as is done in [26, 36].

The instantaneous firing rate at time  $t$  for the  $i^{\text{th}}$  neuron in the network, subject to an external driving rate  $\nu_i(t)$ , is given by

$$\lambda_i(t) = \nu_i(t) + \sum_j \sum_{t_j^s} J_{ij}(t) \epsilon(t - t_j^s - d_{ij}^a). \quad (2)$$

where

$$\epsilon(t) = H(t) \frac{1}{\tau} e^{-t/\tau}$$

is a synaptic filter with  $H$  denoting the Heaviside function and  $\tau$  a synaptic time constant which we set to  $\tau = 5\text{ms}$ . This filter is normalized ( $\int_{-\infty}^{\infty} \epsilon(t) dt = 1$ ) and causal ( $\epsilon(t) = 0$  for  $t < 0$ ) so that only past spikes influence spiking probabilities at  $t$ . Finally, we introduce axonal delays, independently sampled for each synapse from a uniform distribution  $d_{ij}^a \sim \bar{d}^a \pm \sigma^a$  with  $\bar{d}^a = 3\text{ ms}$  and  $\sigma^a = 1$ , a range consistent with experimental studies (see e.g. [28, 29]). We simulate (2) numerically by discretizing time in small bins of size  $\delta t$  and independently drawing spikes for neuron  $i$  in bin  $[t, t + \delta t]$  according to the probability  $\lambda_i(t) \delta t$  (see *Simulation details* below).

**Plastic synapses.** For each pair of pre- and post-synaptic spike times  $t_{pre}$  (from neuron  $j$ ) and  $t_{post}$  (from neuron  $i$ ), the synapse  $J_{ij}$ , if present, will be changed according to a STDP rule  $W$  defined below. This rule defines increments that are added to existing synaptic weights, which are updated every time a new spike is fired. We describe this update procedure below (see Equation (6)). For an inter-spike-interval  $\Delta t = t_{pre} - t_{post}$  at a synapse

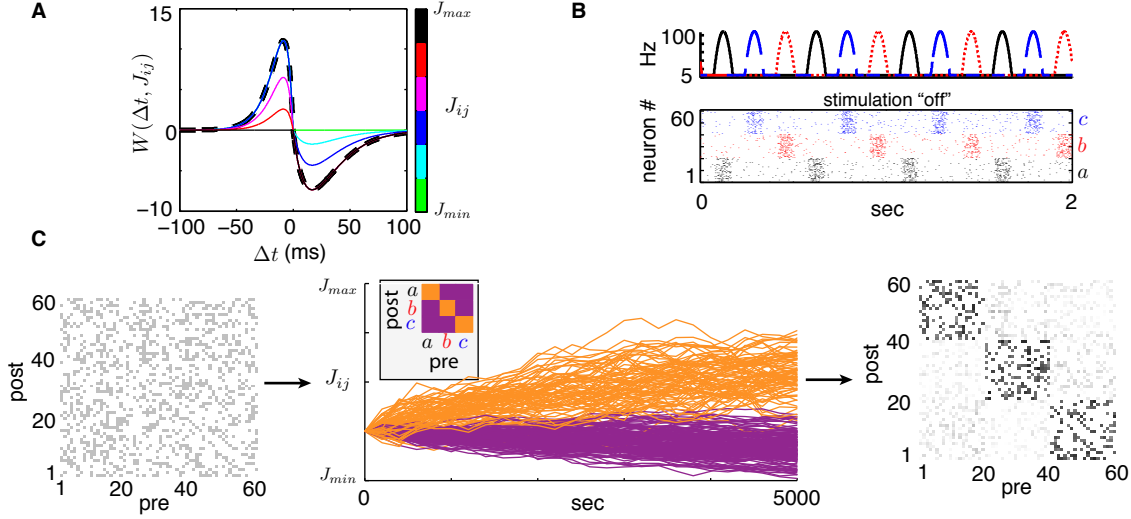


FIG. 7. (A) STDP function  $W(\Delta t, J_{ij})$ . Dashed black line indicates the  $W^+$  and  $W^-$  functions. Coloured lines indicate the full, weight-dependent rule  $W$  for different synaptic weight values  $J_{ij}$  shown in the color bar. (B) Top: Periodic external rates delivered to each neural group (as in Figure 1). Middle: sample spike raster plot from a simulation subject to rates plotted above with sparse, homogeneous connectivity  $J(0)$  ( $N = 60$ ). (C) Simulated synaptic evolution  $J(t)$  with  $J_{min} = 0$ ,  $J_{max} = 0.1$  and  $\eta = 10^{-8}$  under baseline conditions (stimulation “off”) as in B. Left: initial connectivity matrix with connectivity probability  $p = 0.3$  and  $J_{ij}(0) = J_{max}/4$  (grayscale: white =  $J_{min}$ , black =  $J_{max}$ ). Middle: 10 evolving weights  $J(t)$  randomly selected from each of 9 possible pre- and post-synaptic group combinations. Color indicates whether a synapse connects neurons within the same group (orange) or across different groups (purple). Right: synaptic matrix  $J(t)$  at the end of the simulation, once group-averaged weights have reached steady-states.

with weight  $J_{ij}$ , this increment is given by:

$$W(\Delta t, J_{ij}) = \begin{cases} f^+(J_{ij})W^+(\Delta t) & \Delta t < 0 \\ -f^-(J_{ij})W^-(\Delta t) & \Delta t > 0 \end{cases} \quad (3)$$

where  $\Delta t = t_{pre} - t_{post}$ . This is a *multiplicative* STDP rule since it is a product of a weight-dependent term  $f^\pm(J_{ij})$  and a spike-time-dependent term  $W^\pm(\Delta t)$ . Figure 7 A shows a plot of  $W(\Delta t, J_{ij})$  as a function of  $\Delta t$  for various values of  $J_{ij}$ .

The functions  $W^+(\Delta t)$  and  $W^-(\Delta t)$  describe the potentiation and depression of a synapse, respectively, as a function of spike-timing. We assume Hebbian plasticity and require that these functions be both positive and decay to zero as  $|\Delta t|$  grows, so that a pre-synaptic spike preceding a post-synaptic one leads to potentiation while the opposite leads to depression, in accordance with experimental estimates for excitatory STDP [2, 6, 28]. Many functions  $W^\pm$  may lead to similar results; here we choose



$$\begin{aligned}
W^+(\Delta t) &= A^+ \frac{|\Delta t|}{\tau^+} \exp(-|\Delta t|/\tau^+) \\
W^-(\Delta t) &= A^- \frac{|\Delta t|}{\tau^-} \exp(-|\Delta t|/\tau^-)
\end{aligned} \tag{4}$$

as in [17, 24] and others where  $A^+$  and  $A^-$  indicate maximal increments and  $\tau^+$  and  $\tau^-$  are decay time constants. Experimental evidence suggests that  $W^+$  and  $W^-$  are not identical so that STDP is not symmetric, with the maximum potentiation increment often being larger than its depression counterpart ( $A^+ > A^-$ ), and decay lengths following opposite trends ( $\tau^- > \tau^+$ ) [2]. As in [17, 24], respecting generally accepted ranges from various animal models [2, 3, 6, 53], we set:  $A^+ = 30$ ,  $A^- = 20$ ,  $\tau^+ = 8.5$  ms,  $\tau^- = 17$  ms.

The terms  $f^+(J_{ij})$  and  $f^-(J_{ij})$  control the magnitude of plastic increments as a function of synaptic strength and decay to zero as  $J_{ij}$  approaches set bounds  $J_{max}$  and  $J_{min}$  respectively. Thus, synaptic weights can only asymptotically approach these bounds under Equation (3). Here, we choose these functions to be

$$\begin{aligned}
f^+(J_{ij}) &= \left(1 - \frac{J_{ij}}{J_{max}}\right)^\gamma \\
f^-(J_{ij}) &= \left(\frac{J_{ij}}{J_{max}}\right)^\gamma
\end{aligned} \tag{5}$$

where the exponent  $\gamma$  controls the strength of the weight dependence. We set  $\gamma = 0.1$ ,  $J_{min} = 0$ ,  $J_{max} = 0.1$  for our simulations and discuss maximal synaptic strengths in the *Theory* section below. This weight-dependent mechanism reflects the fact that synapses cannot be infinitely strong and that as a synapse potentiates, it becomes harder to potentiate it further as resources to do so deplete (e.g. availability of neurotransmitters and vesicles, etc.). This is consistent with experimental findings [3, 53]. Rather than introducing unrealistic rate-dependent plasticity terms and/or imposing periodic rescaling of synapses, as is often done in the modelling literature, the use of weight-dependent terms in Equation (3) will enable stable connectivity equilibria, akin to homeostasis, while remaining biologically plausible.

In addition to axonal delays  $d_{ij}^a$  introduced earlier, we add dendritic delays independently sampled from a uniform distribution  $d_{ij}^d \sim \bar{d}^d \pm \sigma^d$  with  $\bar{d}^d = 2$  ms and  $\sigma^d = 1$  as an estimate of physiological values (see e.g. [28]). These account for the time for a spike in neuron  $i$  to travel back up dendrites and trigger STDP effects for its synapses. Therefore, for every pair of spike times  $(t_i^s, t_j^s)$ , the synapse  $J_{ij}$  will be updated at time  $t^* = \max(t_i^s + d_{ij}^d, t_j^s + d_{ij}^a)$ :

$$J_{ij}(t^*) \rightarrow J_{ij}(t^*) + \eta W([t_j^s + d_{ij}^a] - [t_i^s + d_{ij}^d], J_{ij}(t^*)) \tag{6}$$

where  $\eta \ll 1$  scales the overall STDP learning rate and is set to  $\eta = 10^{-8}$  for simulations. It is chosen by directly comparing our model's predicted timescale for synaptic changes and the BBCI experiment we aim to capture [22] (see also *Results*).

### Functional groups and spike-triggered stimulation

As illustrated in Figure 1 A and outlined in *Results*, we separate our network in three equally sized groups ( $a$ ,  $b$  and  $c$ ) in which neurons receive the same external drive  $\nu_\alpha(t)$  ( $\alpha = a, b$  or  $c$ ). These groups are indexed in order so that neurons  $i = 1, \dots, N/3$  are in group  $a$ , etc. For exposition, Figure 7 C shows time-traces of the synaptic weights  $J(t)$ , with  $J(0)$  randomly initiated as a sparse matrix with homogeneous weights  $J_{init} = J_{max}/4$ . We can already see some connectivity structure emerge from network activity, with synapses that connect neurons within the same group (in orange) gradually increasing and synapses connecting neurons of different groups (in purple) decreasing.

To mimic the spike-triggered stimulation experiments in [22], we assign group  $a$  to be the “Recording” site, group  $b$  to be the “Stimulation” site and  $c$  to be the “Control” site. Neuron  $i = 1$  from group  $a$  is the “recorded” neuron, whose spikes trigger stimulation of every neuron in group  $b$  after a stimulation delay  $d^\dagger$ . Stimulation of subsets of  $b$  is also considered in *Results*. Thus, we impose the following interactions:

$$\begin{cases} \lambda_i^\dagger(t) = \nu_b(t) + \sum_j J_{ij}(t) \sum_{t_j^s} \epsilon(t - t_j^s - d_{ij}^a) + \sum_{t_1^s} \delta(t - t_1^s - d^\dagger) & ; \quad i \in b \\ \lambda_i^\dagger(t) = \nu_{\alpha_i}(t) + \sum_j J_{ij}(t) \sum_{t_j^s} \epsilon(t - t_j^s - d_{ij}^\alpha) & ; \quad i \notin b \end{cases} \quad (7)$$

where “ $\dagger$ ” indicates the presence of stimulation. The bottom panel of Figure 1 B shows the simulated spiking activity of a network with this stimulation protocol turned “on” and a delay  $d^\dagger = 10$  ms. Further effects of spike-triggered stimulation conditioning are described in *Results*.

### Simulation details

Spiking network simulations were implemented using Python and MATLAB programming languages with the Mersenne Twister algorithm [30] for pseudo-random number generation. A simulation of system (2) is performed by discretizing time in bins of width  $\delta t = 0.1$  ms and pseudo-randomly drawing spikes for each neuron  $i$  according to the probability  $\lambda_i(t) * \delta t$ . For each new spike drawn, all affected synapses are updated according to the summation of rule (6) for all temporally-filtered preceding spikes.

For system (7), in the presence of spike-triggered stimulation, every time a spike from neuron 1 is drawn, the spiking probability of every neuron in group  $b$  is artificially set to 1, after a delay  $d^\dagger$ . Unless otherwise noted, the parameters used for all spiking simulations are listed in Table I.

## Theory

Our goal is to derive an analytical expression for the dynamics of the synaptic matrix  $J(t)$ , to allow us to predict the timescale of plastic changes and relative equilibria. Equations (2) and (6) show that changes of the synapses stored in  $J(t)$ , modulated by the spiking activity of the network, depend on the external rates  $\nu(t)$  and the synaptic matrix  $J(t)$  itself. This leads to a self-consistent relationship between  $J(t)$  dynamics and network spiking activity. Earlier work [13–17] describes a framework to effectively decouple these interactions, which we follow and adapt to our needs.

Parameter	Value	Description
$N$	60	network size (simulations)
$p$	0.3	connection probability
$\tau$	5 (ms)	synaptic time-constant
$\bar{d}^a$	3 (ms)	mean axonal delay
$\sigma^a$	1 (ms)	radius of axonal delay distribution
$\bar{d}^d$	2 (ms)	mean dendritic delay
$\sigma^d$	1 (ms)	radius of dendritic delay distribution
$A^+$	30	STDP potentiation scaling
$A^-$	20	STDP depression scaling
$\tau^+$	8.5 (ms)	STDP potentiation time-constant
$\tau^-$	17 (ms)	STDP depression time-constant
$J_{min}$	0	synaptic strength lower bound
$J_{max}$	0.1	synaptic strength upper bound
$\gamma$	0.1	STDP weight-dependence exponent
$\eta$	$10^{-8}$	STDP plasticity rate
$T$	2 (sec)	plasticity epoch duration
$d^\dagger$	10 (ms)	spike-triggered stimulation delay

TABLE I. Spiking network simulation parameters

The key idea, originally proposed in [24], is to use a separation of timescales, assuming that synapses are locally constant over time intervals  $[t^1, t^2]$ , on which the network has stable and stationary spiking statistics. It follows that accumulated increments for the synapse  $J_{ij}$  over that interval, denoted  $\Delta J_{ij} = J_{ij}(t^2) - J_{ij}(t^1)$ , is given by the sum of all discrete plastic “steps” due to spike-time pairs  $t_i^s, t_j^s \in [t^1, t^2]$  arriving at the  $ij$ -synapse (ignoring delays for clarity):

$$\Delta J_{ij} = \eta \sum_{t^1 \leq t_i^s, t_j^s \leq t^2} W(t_j^s - t_i^s, J_{ij}(t^1)). \quad (8)$$

An equivalent formulation of Equation (8) uses the density of inter-spike-intervals  $u = t_j^s - t_i^s$  over the interval  $[t^1, t^2]$ . To this end, let  $C_{ij}(u)$  be the count of spike pairs separated by  $u$  occurring over the interval  $[t^1, t^2]$ , which is the (non-normalized) cross-correlation between neuron  $j$  and neuron  $i$ . Then,

$$\Delta J_{ij} = \eta \int_{-\infty}^{\infty} du C_{ij}(u) W(u, J_{ij}(t^1)). \quad (9)$$

Equation (9) outlines the basic mechanism governing the evolution of the synaptic matrix  $J(t)$  over consecutive time intervals. The challenge is to express  $C_{ij}(u)$  –itself nontrivially dependent on  $J(t)$  and external rates  $\nu(t)$  – in a closed form, so that Equation (9) can be iterated by updating  $J$  at each step.

The authors of [17] present detailed descriptions of this iteration process, and of ways to estimate the cross-correlations  $C_{ij}(u)$ . However, a number of limitations of this original derivation prevent us from using it directly: (i) The assumption that external rates  $\nu_i(t)$

and  $\nu_j(t)$  are  $\delta$ -correlated; we require arbitrary cross-correlation functions to fit the model to experiments. (ii) It is unclear how to include artificial spike-triggered stimulation induced by the BBICI; this introduces non-trivial statistical dependencies between groups and single neurons.

In the following, still closely following [17], we present a theoretical derivation that addresses these issues. First, we describe parametric constraints that ensure that spiking activity remains stable and does not run away in a self-excitation cascade. Second, we formally describe the timescale separation argument outlined above, followed by estimates for network cross-correlations, both for normal activity and in the presence of spike-triggered stimulation. Finally, we formulate and analyze a dynamical system for averaged synaptic dynamics.

### *Network stability*

Here, we find constraints on synaptic weight bounds to ensure that spiking activity does not grow unbounded due to run-away self-excitation.

To see how this happens, assume that  $J$  is fixed and external rates are constant in time:  $\nu(t) \equiv \nu^0$ . It follows that the mean network rates can be approximated by the following expansion:

$$\bar{\lambda} \sim \nu^0 + J\nu^0 + J^2\nu^0 + \cdots + J^n\nu^0 + \mathcal{O}(J^{n+1}). \quad (10)$$

If the terms in this expansion do not decay fast enough, the average activity rates  $\bar{\lambda}$  quickly diverge. However, it is easy to show that this can be prevented if we require the eigenvalues of  $J$  to remain bounded. For more details about precise stability conditions, see [17] and references therein. For the remainder of this paper, we assume the sufficient, but not necessary condition that all eigenvalues of  $J(t)$  remain within the complex unit circle at all times, and adjust the bounds  $J_{min}$  and  $J_{max}$  accordingly. This means that  $J^n \rightarrow 0$  as  $n \rightarrow \infty$ , and that for any external rates  $\nu(t)$  with stationary statistics, mean rates  $\bar{\lambda}$  are well defined.

### *Separation of timescales*

We now present mathematical arguments in support of the simplifying assumption that synaptic matrices remain constant on short time intervals, since synaptic changes occur at slower timescales than spiking activity. In turn, this leads to the formulation of well-defined cross-correlation quantities for network activity. The arguments presented here closely follow those originally presented in [24] and used throughout [17].

The timescale on which impactful plastic changes occur is much longer than the spiking activity timescale, a fact imposed by  $\eta \ll 1$  which implies that only tiny plastic increments can occur for each pair of pre- and post-synaptic spikes. This allows to effectively separate the spiking dynamics from the synaptic dynamics by assuming that over a reasonably long time period of length  $T$  (on the order of a few seconds), synaptic weights  $J_{ij}(t)$  are approximately constant ( $\approx J_{ij}$ ). As a result, it is possible to derive dynamic equations for the synaptic connectivity matrix  $J$  on the timescale given by  $T$ . Unless otherwise noted, we set  $T = 2$  seconds for the remainder of this paper.

As defined in Equation (2),  $\lambda_i(t)$  gives the instantaneous spiking probability of neuron  $i$  at time  $t$ . This probability depends both on the external rate  $\nu_i(t)$  and on past spikes from

other neurons in the network, which are themselves random processes governed by  $\lambda_j(s)$  for  $s < t$ . In what follows, we consider the ensemble average over all these processes to get the expected rate  $\langle \lambda_i(t) \rangle$ . For the sake of convenience, we drop the brackets and write

$$\lambda_i(t) = \nu_i(t) + \sum_j J_{ij}(\lambda_j * \epsilon_{ij}^a)(t) \quad (11)$$

where we define the delayed convolution as

$$(\lambda_j * \epsilon_{ij}^a)(t) \equiv \int_0^t ds \lambda_j(s) \epsilon(t - d_{ij}^a - s).$$

Equation (11) is defined over plasticity epochs  $[t - T, t]$  where  $J$  is kept constant. At the end of this epoch, we compute the changes in  $J$ .

The density count of spike pairs separated by  $u = t_j^s - t_i^s$  over the epoch  $[t - T, t]$  is given by the (non-normalized) cross-correlation between neurons  $i$  and  $j$ :  $C_{ij}(u; t) = \langle \lambda_i(s) \lambda_j(s + u) \rangle_t \equiv \int_{t-T}^t ds \lambda_i(s) \lambda_j(s + u)$ . In matrix form, we get

$$C(u; t) = \langle \lambda(s) \lambda^\top(s + u) \rangle_t \quad (12)$$

with  $\lambda(t) = (\lambda_1(t), \dots, \lambda_N(t))^\top$  and “ $\top$ ” denoting matrix transposition. For “post/pre” neurons  $(i, j)$ , the expected number of spike pairs separated by  $\Delta t \in [u_1, u_2]$  in epoch  $[t - T, t]$  is given by  $\int_{u_1}^{u_2} du C_{ij}(u; t)$ . It follows that the expected incremental change for the synapse  $J_{ij}$ , after the epoch  $[t - T, t]$ , is given by

$$\langle \Delta J_{ij}(t) \rangle = \eta \int_{-\infty}^{\infty} du C_{ij}(u + d_{ij}^d; t) W(u - d_{ij}^a, J_{ij}(t - T)). \quad (13)$$

Expression (13) illustrates that all relevant information needed to predict expected synaptic changes is contained in the network’s cross-correlation  $C(u; t)$ .

### *Cross-correlations for normal network activity*

We now derive estimates for cross-correlations of network spiking activity  $C(u)$  in terms of  $J$  and the statistics of the external driving rates  $\nu(t)$ , so that network statistics can be expressed without implicit references to spiking probabilities  $\lambda(t)$ . While a similar quantity is used in [17], we introduce a new expansion-based approach that generalizes to arbitrary external statistics.

Consider the (non-normalized) cross-correlation between external rates:

$$\hat{C}(u; t) = \langle \nu(s) \nu^\top(s + u) \rangle_t$$

defined as in Equation (12). We assume that such “external” cross-correlations are stationary in time and therefore omit the dependence on  $t$ :  $\hat{C}(u)$ . For our example rates  $\nu(t)$  from Figure 1 C,  $\hat{C}(u)$  is plotted in Figure 8 (thin black line). How will the network react to external drives with these statistics?

We start by replacing every delay in the network by its mean. As long as the delay distributions are not too wide, this gives accurate results. This enables us to write (11) in

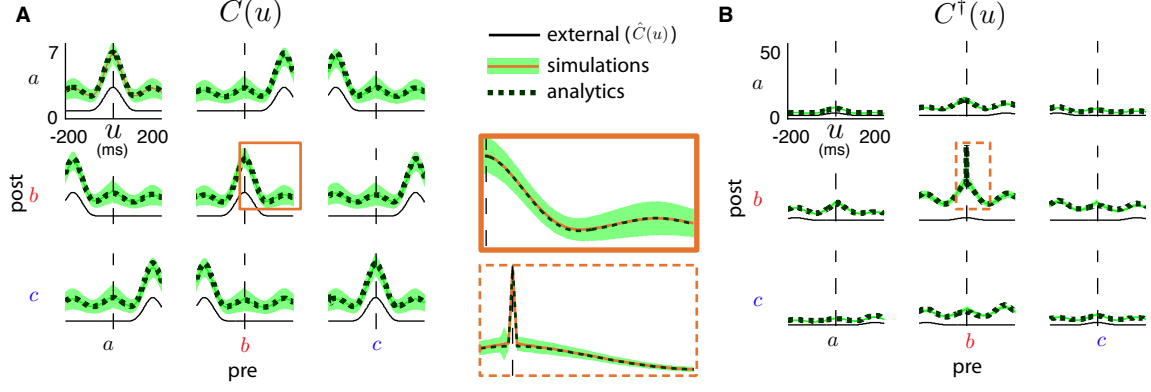


FIG. 8. Correlation functions averaged over all pairs of connected cells in respective group combinations: external rate correlations  $\hat{C}(u)$  (solid black); estimates for network correlations from full spiking network simulations (orange line with shaded green area showing standard deviation of 5 seconds batched estimates) and analytical estimates of order 4 (dashed green). (A) For network under normal activity. (B) For network under spike-triggered stimulation ( $\dagger$ ). Boxes in the middle show zooms from  $C^{\dagger}(u)$  from A and B.

matrix form:

$$\lambda(t) = \nu(t) + J(\lambda * \epsilon^a)(t) \quad (14)$$

where  $\epsilon^a(t)$  is shorthand for  $\epsilon(t - \bar{d}^a)$ . Ideally, one would like to isolate  $\lambda$  from Equation (14) but this is difficult because of the convolution with the synaptic kernel  $\epsilon$ . In [13–17], the authors circumvent this problem by performing calculations in Laplace space, thus transforming the convolution into a product and enabling a self-consistent equation for  $\lambda(t)$ . However, this required important assumptions to be made about the shape of external cross-correlations  $\hat{C}(u)$  so that inverse transforms remain tractable. In our case, we need  $\hat{C}(u)$  to remain as general as possible which makes solving a self-consistent equation exactly challenging.

To circumvent this problem we rely on the stability condition that  $J$  must have eigenvalues within the unit circle, and use this to produce expansions that we truncate at higher powers of  $J$ . We start by substituting Equation (14) into the expression (12) for  $C(u)$ :

$$C(u) = \hat{C}(u) + \langle \nu(s)(\lambda * \epsilon^a)^\top(s+u) \rangle J^\top + J \langle (\lambda * \epsilon^a)(s) \nu^\top(s+u) \rangle + JC(u)J^\top. \quad (15)$$

To simplify further, we replace the cross-correlation terms involving convolved rates with delayed correlations

$$\langle (\lambda * \epsilon^a)(s) \nu^\top(s+u) \rangle \simeq \langle \lambda(s - \bar{d}^a) \nu^\top(s+u) \rangle$$

which is justified for  $T$  large enough and  $\nu(t)$  fluctuating on longer timescales than  $\tau$  ( $\epsilon$ 's time-constant). In turn, we can further expand this expression to obtain:

$$\langle \lambda(s - \bar{d}^a) \nu^\top(s+u) \rangle = \sum_{k=0}^n J^k \hat{C}(u + (k-1)\bar{d}^a) + \mathcal{O}(J^{(n+1)}) \quad (16)$$

with  $\mathcal{O}(J^{(n+1)}) \rightarrow 0$  as  $n \rightarrow \infty$ . A similar expression can be derived for  $\langle \nu(s)(\lambda * \epsilon^a)^\top(s+u) \rangle$ .



Combining Equations (15) and (16), we obtain the following truncated,  $n^{\text{th}}$ -order expansion formula:

$$C(u) \simeq \sum_{k=0}^n \sum_{\substack{r,l \geq 0 \\ r+l=k}} J^r \hat{C}(u + (r-l)\bar{d}^a) J^{\top l} \quad (17)$$

that provides an algorithm to estimate self-consistent network cross-correlations. Figure 8 A shows  $C(u)$  averaged over functional groups, both estimated from spiking simulations and obtained from Equation (17) truncated to fourth order; they agree nicely.

### *Cross-correlations for spike-triggered stimulated networks*

We next adapt our expansion-based approach to derive an expression for  $C^\dagger(u)$ : the network's cross-correlation under spike-triggered stimulation. Novel challenges arise since complex statistical dependencies emerge from triggering artificial stimulation on single neuron spikes, and are beyond the scope of the original framework described in [17].

Recall that after a delay  $d^\dagger$ , the entirety of neural group  $b$  is artificially stimulated and forced to spike whenever neuron 1 from group  $a$  fires a spike (see Equation (7)). This is easily implementable in numerical simulations but raises a conceptual difficulty when attempting to transform it into a consistent expression for the densities  $\lambda^\dagger(t)$ . This is because a particular realization of spike times from neuron 1 must be fixed across all neurons  $i$  in group  $b$ . This is akin to a Hawkes process [20] and demands special corrections to be made to  $C^\dagger(u)$  estimates. We revisit these corrections; for now, we follow the same logic as for Equation (14) and write

$$\lambda^\dagger(t) = \nu(t) + A\lambda(t - d^\dagger) + J(\lambda * \epsilon^a)(t) \quad (18)$$

where  $A_{j1} = 1$  for  $j \in b = \{N/3 + 1, \dots, 2N/3\}$ . The matrix  $A$  adds a copy of the spiking probability from neuron 1 to all neurons in group  $b$ , after a delay  $d^\dagger$ .

We replace the synaptic filter convolution with the delayed rate, as done in (16), directly in Equation (18), while keeping in mind these terms will be combined later into cross-correlation terms. We get

$$\lambda^\dagger(t) = \sum_{k=0}^n \sum_{\substack{r,s \geq 0 \\ r+l=k}} \Pi(A^r, J^l) \nu(t - rd^\dagger - l\bar{d}^a) + \mathcal{O}(J^n) \quad (19)$$

where  $\Pi(A^r, J^l)$  denotes the sum of all possible ordered permutations of  $r$  times  $A$  and  $l$  times  $J$  (e.g.  $\Pi(A^1, J^2) = AJ^2 + J^2A + JAJ$ ). In turn, this leads to the following truncated estimate:

$$C^\dagger(u) \simeq \sum_{k=0}^n \sum_{\substack{r,l,r',l' \geq 0 \\ r+l+r'+l'=k}} \Pi(A^r, J^l) \left[ \hat{C}(u + (r-r')d^\dagger + (l-l')\bar{d}^a) \Pi(A^{\top r'}, J^{\top l'}) \right] + R(u; \bar{\lambda}^\dagger) \quad (20)$$

where  $R(u; \bar{\lambda}^\dagger)$  is a correction term defined below. As delays accumulate and since the matrices  $A$  and  $J$  do not necessarily commute, there is no simple way to concisely express Equation (20). Nevertheless, its simple combinatorial nature enables a relatively straightforward algorithmic implementation. Moreover, while comparing estimates to numerical simulations

with current parameters, we find very good agreement with third- or fourth-order estimates.

We still need to account for the fact that a single spike train  $S_1(t) \sim \lambda_1^\dagger(t)$  is copied into group  $b$  rather than independent realizations sharing the same rate  $\lambda_1^\dagger(t)$ . Since the rates  $\lambda^\dagger(t)$  are treated as ensemble averages ( $\langle \lambda^\dagger(t) \rangle$ ) and assuming independence between disjoint time intervals (as is the case for Poisson processes), it follows that  $\langle \lambda_i^\dagger(s) S_1(s+u) \rangle = \langle \lambda_i^\dagger(s) \lambda_1^\dagger(s+u) \rangle$  if  $u \neq 0$ . To account for the fact that all neurons in group  $b$  are driven by a single spike copy, the following modifications are made via the correction term  $R(u; \bar{\lambda}^\dagger)$  in Equation (20):

$$\begin{aligned} R_{ij}(u; \bar{\lambda}^\dagger) &= T \bar{\lambda}_1^\dagger \delta(u) \\ R_{i1}(u; \bar{\lambda}^\dagger) &= T \bar{\lambda}_1^\dagger \delta(u - d^\dagger) \\ R_{1i}(u; \bar{\lambda}^\dagger) &= T \bar{\lambda}_1^\dagger \delta(u + d^\dagger). \end{aligned} \tag{21}$$

where  $i, j \in b$  and  $\bar{\lambda}_1^\dagger$  is defined as in Equation (10) with  $\nu^0 = \frac{1}{T} \int_{t-T}^t \nu(s) ds$  (the mean over the epoch) using the same expansion truncation order  $n$  used in the cross-correlation estimate. Figure 8 B shows agreement between group-averaged  $C^\dagger(u)$  derived from Equations (19), (21) and estimated from spiking numerical simulations.

### *Dynamical system for averaged synaptic strengths*

Using the estimates for network cross-correlations with or without BBICI stimulation,  $C^{(\dagger)}(u)$ , we derive a dynamical system for synaptic weights averaged over neural groups. This dynamical system has a temporal resolution of  $T$ , as it describes the evolution of synapses subject to STDP every  $T$ -step. Stability arguments are also presented, describing conditions under which one can expect stable, mean synaptic equilibria. Here, we closely follow [17] although stability arguments differ slightly due to the differences in our model presented above.

In order to use the correlation estimates to predict synaptic changes, we begin by restricting ourselves to group-averaged quantities. This greatly simplifies calculations and gives remarkably precise estimates for networks where  $N$  is large enough. Consider the  $3 \times 3$  matrix  $\bar{J}$  whose entries  $\bar{J}_{\alpha\beta}$  represent the mean strength of non-zero synapses from a neuron in group  $\beta$  to a neuron in group  $\alpha$ . Recall that  $p$  is the probability that any two cells are connected which means that  $\frac{pN}{3} \bar{J}_{\alpha\beta}$  is the mean strength of total synaptic inputs from neurons in group  $\beta$  to a single neuron in group  $\alpha$ .

We follow derivations outlined in the previous section to obtain estimates for the group-averaged network cross-correlations. For tractability, we explicitly index these estimates by the (averaged) synaptic matrix  $\bar{J}$  and external cross-correlation  $\hat{C}(u)$  used in their derivation:

$\bar{C}_{\bar{J},\hat{C}}(u)$  and  $\bar{C}_{\bar{J},\hat{C}}^\dagger(u)$ . These are now  $u$ -dependent  $3 \times 3$  matrices obtained by using

$$\begin{aligned} J &= \frac{pN}{3} \bar{J} \\ \hat{C}(u) &= \langle [\nu_a, \nu_b, \nu_c](s) [\nu_a, \nu_b, \nu_c]^\top(s+u) \rangle \\ A &= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

in equations (17) and (20). Furthermore, the correction term  $R(u; \bar{\lambda}^\dagger)$  in Equation (20) following this group average is replaced by its  $3 \times 3$  averaged counterpart  $\bar{R}(u; \bar{\lambda}^\dagger)$  which is zero everywhere except:

$$\begin{aligned} \bar{R}_{bb}(u; \bar{\lambda}^\dagger) &= T \bar{\lambda}_a^\dagger \delta(u) \\ \bar{R}_{ba}(u; \bar{\lambda}^\dagger) &= \frac{3}{N} T \bar{\lambda}_a^\dagger \delta(u - d^\dagger) \\ \bar{R}_{ab}(u; \bar{\lambda}^\dagger) &= \frac{3}{N} T \bar{\lambda}_a^\dagger \delta(u + d^\dagger). \end{aligned} \tag{22}$$

Note that for large  $N$ , the modification for interactions between groups  $a$  and  $b$  vanish, but interactions within group  $b$  remain independent of  $N$ , since each cell in  $b$  shares the same artificially introduced spikes (see Figure 8 B).

Having a closed expression for  $\bar{C}_{\bar{J},\hat{C}}^{(\dagger)}(u)$ , we define the functional

$$F^{(\dagger)}(M, \bar{J}; \hat{C}) \equiv \int_{-\infty}^{\infty} du \bar{C}_{\bar{J},\hat{C}}^{(\dagger)}(u + \bar{d}^d) \circ W(u - \bar{d}^a, M) \tag{23}$$

where “ $\circ$ ” designates an element-wise (Hadamard) product and  $\bar{C}_{\bar{J},\hat{C}}^{(\dagger)}$  refers to the estimated network correlations using mean matrix  $\bar{J}$  and  $\hat{C}(u)$  with and without spike-triggered stimulation ( $\dagger$ ) respectively. From Equation (13),  $F$  gives the expected average plastic increments for synaptic matrix  $M$  (also  $3 \times 3$ ), over the epoch  $[t - T, t]$ . Figure 9 A shows  $F_{\alpha\beta}(M, \bar{J}; \hat{C})$  for  $\bar{J} = \bar{J}(0)$ , a random sparse homogeneous matrix as in Figure 7, and  $\hat{C}(u)$  as in Figure 8. It follows that  $F$  defines a discrete, state-dependent dynamical system on the space of  $3 \times 3$  matrices describing the evolution of our network’s synapses from one activity epoch to the next:

$$\bar{J}(t + T) = \bar{J}(t) + \eta F(\bar{J}(t), \bar{J}(t); \hat{C}). \tag{24}$$

For any initial matrix  $\bar{J}(0)$ , iterating Equation (24) gives the synaptic evolution of  $\bar{J}(t)$  at a temporal resolution  $T$ .

For stationary external statistics (i.e.  $\hat{C}(u)$  does not depend on  $t$ ), a fixed point of the system described by (24) is a matrix  $\bar{J}^*$  such that  $F(\bar{J}^*, \bar{J}^*; \hat{C}) = 0$  and defines an averaged synaptic equilibrium matrix. Furthermore, an equilibrium is said to be *stable* if nearby initial states ( $\bar{J}(0)$ ) converge toward it. We are interested in the existence and stability of equilibria as they represent homeostatic steady states. However, classical tools of dynamical systems theory such as linearization prove to be inadequate because  $F$  is an integro-functional. Nevertheless, some basic geometric observation can provide valuable insights.

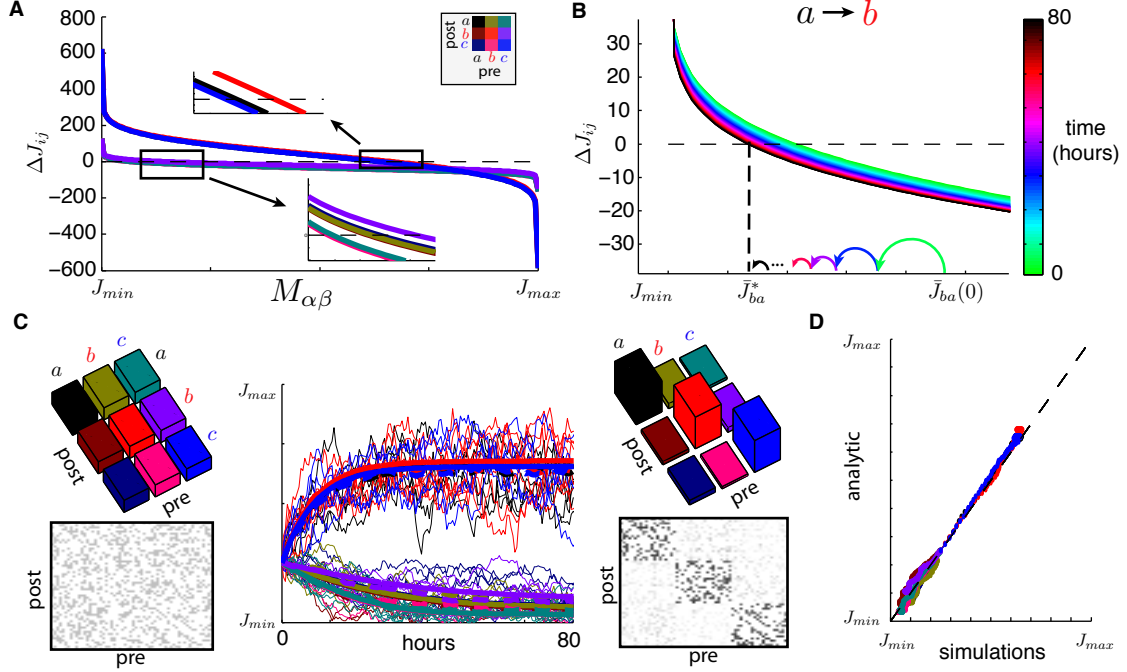


FIG. 9. (A) Plot of  $F(M_{\alpha\beta}, \bar{J}(0), \hat{C}(u))$  (with  $\alpha, \beta \in \{a, b, c\}$ ) where  $J(0)$  is sparse and homogeneous (see C) and  $\hat{C}(u)$  is as in Figure 8. Colors represent the pair of pre- and post-synaptic groups  $\alpha\beta$  as in the upper left key. (B) Plot of  $F(M_{ba}, \bar{J}_{ba}(t), \hat{C}(u))$  as  $\bar{J}(t)$  evolves over time indicated by the color scheme. Curved arrows illustrate the iteration scheme  $\bar{J}_{ba}(t+T) = \bar{J}_{ba}(t) + F_{ba}(\bar{J}(t), \bar{J}(t), \hat{C})$  and culminate at the equilibrium  $\bar{J}_{ba}^*$ . (C) Middle: evolution of synaptic weights over time, color-coded as in A. Thick solid lines show analytical predictions of order 4, thick dotted lines show averages from simulations and thin lines show examples of individual synapses  $J_{ij}(t)$ . On either sides: plots of  $\bar{J}(t)$  and  $J(t)$  used for simulations for  $t = 0$  (left) and  $t = 80$  hours (right). (D) Scatter plot of  $\bar{J}_{\alpha\beta}(t)$  and corresponding averages of simulated  $J(t)$  over all time points plotted in C. For both C and D,  $N = 60$ ,  $p = 0.3$ ,  $J_{min} = 0$ ,  $J_{max} = 0.1$ ,  $\eta = 10^{-8}$  and  $J_{ij}(0) = J_{max}/4$ .

First, notice that  $\bar{C}(u) \geq 0$  since it is defined as the non-normalized cross-correlation. This means that for  $M_{ij}$  close to the bounds  $\bar{J}_{min}$  and  $\bar{J}_{max}$ —where  $W(\Delta t, M_{ij})$  is positive and negative, respectively, for any  $\Delta t$ — $F$  must be positive and negative respectively, for any matrix  $\bar{J}$  and any  $\hat{C}(u)$ . Moreover, from the definition of  $f^\pm(M_{ij})$  in Equation (5), it is straightforward to see that all entries of  $F(M, J; \hat{C})$  are monotonically decreasing with respect to each  $M_{ij}$  over the interval  $[J_{min}, J_{max}]$ . Since  $F$  is clearly continuous in  $M$ , it follows from the mean value theorem that there exists a unique  $M^*(\bar{J})$  for any  $\bar{J}$  such that

$$F(M^*(\bar{J}), \bar{J}, \hat{C}) = 0. \quad (25)$$

This is illustrated in Figure 9 A where each curve intersects the  $\Delta \bar{J}_{\alpha\beta} = 0$  line exactly once. It follows that a matrix  $\bar{J}^*$  is a fixed point of system (24) if and only if it is also a fixed point of  $M^*$ , i.e.,  $\bar{J}^* = M^*(\bar{J}^*)$  in Equation (25). It is not clear that there exists a general condition that guarantees that  $M^*$  admits a fixed point for finite values of  $\eta$ . Nevertheless we argue that in general, one can expect the dynamics of (24) to converge toward a very

small region of matrix space and remain there, as is illustrated in Figure 9 B.

To see why this should be the case, let us assume without loss of generality that at some time  $t$ ,  $\bar{J}_{\alpha\beta}(t) > M_{\alpha\beta}^*(\bar{J}(t))$ . By construction, this means that  $F_{ij}(\bar{J}, \bar{J}; \hat{C}) < 0$  so that  $\bar{J}_{\alpha\beta}(t+T) < \bar{J}_{\alpha\beta}(t)$  (i.e.  $\bar{J}_{\alpha\beta}$  decreases). Assuming very small steps and by the continuity of  $F$ , we can expect  $M^*$  to also show small changes as  $\bar{J}_{\alpha\beta}$  evolves. Thus, the only way the distance between  $\bar{J}_{\alpha\beta}$  and  $M_{\alpha\beta}^*(\bar{J})$  can grow is if  $M_{\alpha\beta}^*(\bar{J}(t))$  also decreases, but this cannot go on forever since  $M_{\alpha\beta}^*$  is bounded. It follows that beyond a certain time,  $|\bar{J}_{\alpha\beta}(t+T) - M_{\alpha\beta}^*(\bar{J}(t+T))| < |\bar{J}_{\alpha\beta}(t) - M_{\alpha\beta}^*(\bar{J}(t))|$  which implies that  $\bar{J}_{\alpha\beta}$  converges toward  $M_{\alpha\beta}^*(\bar{J})$ .

This contraction argument would be sufficient to show the existence of a fixed point  $\bar{J}_{\alpha\beta}^*$  if the function  $M_{\alpha\beta}^*$  depended only on the  $\alpha\beta$  component of  $\bar{J}$ , but this is not the case by construction of the functional  $F$ . Indeed, interactions throughout the network influence the cross-correlation between any two groups  $\alpha$  and  $\beta$  which in turn, influence expected synaptic changes between these groups. However, if synapses converge toward their respective fixed points at a comparable rate, then their contributions to  $M^*$  all attenuate at similar rates, leading to stable steady states. This is what we observe when we numerically iterate system (24).

Figure 9 B illustrates this stability property by showing the “ $ba$ ” coordinate of  $F$  throughout the iteration process (24) along with a cartoon representation of iterates  $\bar{J}_{ba}(t)$ . We can clearly see that not only does the iterate push  $J(t)$  toward a value for which  $F = 0$ , but  $F$  also changes in a monotonic manner so that for small enough steps modulated by  $\eta$ ,  $\bar{J}(t)$  is bound to converge to a stable  $\bar{J}^*$ . The same is true for each coordinate leading to a global  $J^*$ .

Figure 9 C shows both spiking network simulations and averaged synaptic dynamics for a system with normal dynamics (i.e. no stimulation) starting from a sparse homogeneous matrix. Equilibria for averaged synaptic strengths across all functional groups are well captured by the reduced system. Note that while mean synaptic strengths are very stable once an equilibrium is reached, individual synapses are not; they continue to fluctuate as a result of irregular spiking activity and inhomogeneous connectivity.

Moreover, the reduced system (24) also captures the timescale at which group-averaged synapses converge to their respective equilibria since it estimates the expected synaptic changes over plasticity epochs, modulated by the learning rate  $\eta$ . To better appreciate this point, Figure 9 D shows a scatter plot of  $\bar{J}(t)$  from system (24) against spiking simulations for all time points plotted in panel C of the same figure: all points lie close to the diagonal, indicating a tight agreement throughout the convergence process. We will see in the *Results* section that capturing synaptic convergence timescales with our reduced model enables us to calibrate the learning rate  $\eta$  to fit experimental observations.

Finally, note that to obtain estimates for equilibria alone, a complete iteration procedure is not required. Using the contraction argument presented above, it is sufficient to evaluate  $M^*(\bar{J})$  from Equation (25) for any initial matrix  $\bar{J}$ , and repeat the process a few times (i.e.  $M^*(M^*(\dots M^*(\bar{J})))$ ). We find that two or three iterations of this process are enough to converge to the equilibrium  $\bar{J}^*$  within machine precision.

In summary, we showed that it is possible to obtain estimates for the network’s cross-correlation at a temporal resolution defined by plasticity epochs  $[t - T, t]$ , for both the normal condition ( $C(u; t)$ ) and the spike-triggered condition ( $C^\dagger(u; t)$ ). These estimates can be truncated at any order (i.e. ignoring terms of higher powers of  $J$ ) and depend only on

the cross-correlation of external inputs over the same plasticity epoch ( $\hat{C}(u; t)$ ). We find through comparison with numerical simulations that truncating to third or fourth order is sufficient to capture network statistics almost perfectly as is shown in Figure 8. These estimates can then be used to create a dynamical system capable of estimating both the stable synaptic equilibria that emerge from spiking dynamics and the convergence timescales leading to them, for both the normal and spike-triggered stimulation paradigms.

## Experimental Procedure

In *Results*, we use experimentally obtained cross-correlation functions from spiking activity recorded in a monkey implanted with a Utah array in primary MC and performing an isometric 2D target-tracking task. Spikes from multiple electrodes of the array were recorded during task performance, and cross-correlograms were compiled with a resolution of 2 milliseconds. Recorded activity was modulated during task performance. We assume the neurons whose spikes are recorded from the same electrode (using spike sorting) were close enough to be part of the same neural group, in the context of our model. In contrast, neurons recorded from different electrodes, separated by at least 400  $\mu\text{m}$ , are considered as part of different groups.

## ACKNOWLEDGMENTS

The authors thank Alain Vinet for insightful discussions, and Andrew Bogaard and Andrew Jackson for experimental data. This work was funded in part by the Washington Research Foundation (WRF) UW Institute for Neuroengineering (*GL*), the National Science Foundation (NSF) Center for Sensorimotor Neural Engineering at the University of Washington, NIH grant NS12542 (*EF*), an Allen Distinguished Investigator grant (*ALF*), and CIHR grant MOP-84454 and HFSP RGP0049/2009-C (*JFK* and *NIK*).

- 
- [1] Bazhenov, M., Timofeev, I., Steriade, M., and Sejnowski, T. J. (2002). Model of Thalamocortical Slow-Wave Sleep Oscillations and Transitions to Activated States. The Journal of neuroscience : the official journal of the Society for Neuroscience, 22(19):8691–8704.
  - [2] Bi, G. and Poo, M. (2001). Synaptic modification by correlated activity: Hebb’s postulate revisited. Annual Review of Neuroscience.
  - [3] Bi, G. Q. and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. The Journal of neuroscience : the official journal of the Society for Neuroscience, 18(24):10464–10472.
  - [4] Buonomano, D. V. and Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. Annual review of neuroscience, 21(1):149–186.
  - [5] Burkitt, A. N., Gilson, M., and Van Hemmen, J. L. (2007). Spike-timing-dependent plasticity for neurons with recurrent connections. Biological Cybernetics, 96(5):533–546.
  - [6] Caporale, N. and Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. Annual Review of Neuroscience, 31(1):25–46.



- [7] Carmenta, J. M., Lebedev, M. A., Crist, R. E., O'Doherty, J. E., Santucci, D. M., Dimitrov, D. F., Patil, P. G., Henriquez, C. S., and Nicolelis, M. A. L. (2003). Learning to Control a Brain–Machine Interface for Reaching and Grasping by Primates. PLoS Biology, 1(2):e42.
- [8] Carrillo-Reid, L., Yang, W., Bando, Y., Peterka, D. S., and Yuste, R. (2016). Imprinting and recalling cortical ensembles. Science (New York, NY), 353(6300):691–694.
- [9] Cheney, P. D. and Fetz, E. E. (1980). Postspike facilitation of forelimb muscle activity by primate corticomotoneuronal cells. Journal of Neurophysiology, 44(4):751–772.
- [10] Churchland, M. M. and Shenoy, K. V. (2007). Temporal Complexity and Heterogeneity of Single-Neuron Activity in Premotor and Motor Cortex. Journal of Neurophysiology, 97(6):4235–4257.
- [11] Edwardson, M. A., Lucas, T. H., Carey, J. R., and Fetz, E. E. (2013). New modalities of brain stimulation for stroke rehabilitation. Experimental Brain Research, 224(3):335–358.
- [12] Fetz, E. E. (1984). Functional organization of motor and sensory cortex symmetries and parallels. . In Dynamic Aspects of Neocortical Function, pages 453–473. J. Wiley.
- [13] Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and Hemmen, J. L. v. (2009a). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. I. Input selectivity–strengthening correlated input pathways. Biological Cybernetics, 101(2):81–102.
- [14] Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and Hemmen, J. L. v. (2009b). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. II. Input selectivity—symmetry breaking. Biological Cybernetics, 101(2):103–114.
- [15] Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and Hemmen, J. L. v. (2009c). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV. Biological Cybernetics, 101(5-6):427–444.
- [16] Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009d). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks III: Partially connected neurons driven by spontaneous activity. Biological Cybernetics, 101(5-6):411–426.
- [17] Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2010). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks V: self-organization schemes and weight dependence. Biological Cybernetics, 103(5):365–386.
- [18] Guggenmos, D. J., Azin, M., Barbay, S., Mahnken, J. D., Dunham, C., Mohseni, P., and Nudo, R. J. (2013). Restoration of function after brain damage using a neural prosthesis. Proceedings of the National Academy of Sciences of the United States of America, 110(52):21177–21182.
- [19] Güttig, R., Aharonov, R., Rotter, S., and Sompolsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. The Journal of neuroscience : the official journal of the Society for Neuroscience, 23(9):3697–3714.
- [20] Hawkes, A. (1971). Point spectra of some mutually exciting point processes. Journal of the Royal Statistical Society. Series B ( ...).
- [21] Jackson, A., Gee, V., Baker, S., and Lemon, R. (2003). Synchrony between neurons with similar muscle fields in monkey motor cortex. Neuron.
- [22] Jackson, A., Mavoori, J., and Fetz, E. E. (2006). Long-term motor cortex plasticity induced by an electronic neural implant. Nature, 444(7115):56–60.
- [23] Jackson, A., Mavoori, J., and Fetz, E. E. (2007). Correlations Between the Same Motor Cortex Cells and Arm Muscles During a Trained Task, Free Behavior, and Natural Sleep in

- the Macaque Monkey. Journal of Neurophysiology, 97(1):360–374.
- [24] Kempster, R., Gerstner, W., and Van Hemmen, J. (1999). Hebbian learning and spiking neurons. Physical Review E.
  - [25] Ledochowitsch, P., Yazdan-Shahmorad, A., Bouchard, K. E., Diaz-Botia, C., Hanson, T. L., He, J. W., Seybold, B. A., Olivero, E., Phillips, E. A. K., Blanche, T. J., Schreiner, C. E., Hasenstaub, A., Chang, E. F., Sabes, P. N., and Maharbiz, M. M. (2015). Strategies for optical control and simultaneous electrical readout of extended cortical circuits. Journal of neuroscience methods, 256(C):220–231.
  - [26] Lindner, B., Doiron, B., and Longtin, A. (2005). Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback. Physical Review E, 72(6):61919–1179.
  - [27] Lucas, T. and Fetz, E. E. (2013). Myo-cortical crossed feedback reorganizes primate motor cortex output. The Journal of Neuroscience, 33(12):5261–5274.
  - [28] Markram, H., Lübke, J., Frotscher, M., and Roth, A. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. The Journal of Physiology, 500(2):409–440.
  - [29] Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., Kahou, G. A. A., Berger, T. K., Bilgili, A., Buncic, N., Chalimourda, A., Chindemi, G., Courcol, J.-D., Delalandre, F., Delattre, V., Druckmann, S., Dumusc, R., Dynes, J., Eilemann, S., Gal, E., Gevaert, M. E., Ghobril, J.-P., Gidon, A., Graham, J. W., Gupta, A., Haenel, V., Hay, E., Heinis, T., Hernandez, J. B., Hines, M., Kanari, L., Keller, D., Kenyon, J., Khazen, G., Kim, Y., King, J. G., Kisvarday, Z., Kumbhar, P., Lasserre, S., Le Bé, J.-V., Magalhães, B. R. C., Merchán-Pérez, A., Meystre, J., Morrice, B. R., Muller, J., Muñoz-Céspedes, A., Muralidhar, S., Muthurasa, K., Nachbaur, D., Newton, T. H., Nolte, M., Ovcharenko, A., Palacios, J., Pastor, L., Perin, R., Ranjan, R., Riachi, I., Rodríguez, J.-R., Riquelme, J. L., Rössert, C., Sfyarakis, K., Shi, Y., Shillcock, J. C., Silberberg, G., Silva, R., Tauheed, F., Telefont, M., Toledo-Rodriguez, M., Tränkler, T., Van Geit, W., Díaz, J. V., Walker, R., Wang, Y., Zaninetta, S. M., DeFelipe, J., Hill, S. L., Segev, I., and Schürmann, F. (2015). Reconstruction and Simulation of Neocortical Microcircuitry. Cell, 163(2):456–492.
  - [30] Matsumoto, N. and Okada, M. (2004). Impact of deviation from precise balance of spike-timing-dependent plasticity. Neural Networks, 17(7):917–924.
  - [31] Meffin, H., Besson, J., Burkitt, A., and Grayden, D. (2006). Learning the structure of correlated synaptic subgroups using stable and competitive spike-timing-dependent plasticity. Physical Review E.
  - [32] Monfils, M.-H., Plautz, E. J., and Kleim, J. A. (2005). In search of the motor engram: motor map plasticity as a mechanism for encoding motor experience. The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry, 11(5):471–483.
  - [33] Morrison, A., Aertsen, A., and Diesmann, M. (2007). Spike-timing-dependent plasticity in balanced random networks. Neural Computation, 19(6):1437–1467.
  - [34] Mountcastle, V. B. and Powell, T. P. (1959). Neural mechanisms subserving cutaneous sensibility, with special volume 32. Bulletin of the Johns Hopkins Hospital.
  - [35] Nishimura, Y., Perlmutter, S. I., Eaton, R. W., and Fetz, E. E. (2013). Spike-Timing-Dependent Plasticity in Primate Corticospinal Connections Induced during Free Behavior. Neuron, 80(5):1301–1309.

- [36] Ocker, G. K., Litwin-Kumar, A., and Doiron, B. (2015). Self-Organization of Microcircuits in Networks of Spiking Neurons with Plastic Synapses. PLoS Computational Biology, 11(8):e1004458–40.
- [37] Paninski, L., Fellows, M. R., Hatsopoulos, N. G., and Donoghue, J. P. (2004). Spatiotemporal Tuning of Motor Cortical Neurons for Hand Position and Velocity. Journal of Neurophysiology, 91(1):515–532.
- [38] Potter, S. M., El Hady, A., and Fetz, E. E. (2014). Closed-loop neuroscience and neuroengineering. Frontiers in Neural Circuits, 8.
- [39] Rebesco, J. M., Stevenson, I. H., Koerding, K., Solla, S. A., and Miller, L. E. (2010). Rewiring Neural Interactions by Micro-Stimulation. Frontiers in Systems Neuroscience, 4.
- [40] Rubin, J. E., Gerkin, R. C., Bi, G.-q., and Chow, C. C. (2005). Calcium Time Course as a Signal for Spike-Timing-Dependent Plasticity. Journal of Neurophysiology, 93(5):2600–2613.
- [41] Sanes, J. N. and Donoghue, J. P. (2000). Plasticity and primary motor cortex. Annual review of neuroscience.
- [42] Schieber, M. H. (2001). Constraints on Somatotopic Organization in the Primary Motor Cortex. Journal of Neurophysiology, 86(5):2125–2143.
- [43] Seeman, S., Mogen, B. J., Fetz, E. E., and Perlmutter, S. (2015). Paired stimulation induces spike-timing dependent plasticity of neural connections in primate sensorimotor cortex. In Society for Neuroscience Annual Meeting proceedings.
- [44] Silversmith, D. B., Yazdan-Shahmorad, A., and Sabes, P. N. (2016). Targeted cortical reorganization using optogenetics in non-human primates. In Society for Neuroscience Annual Meeting proceedings.
- [45] Smith, M. A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. PLoS Biology, 4(6):e179–9.
- [46] Smith, W. and Fetz, E. E. (2009a). Synaptic interactions between forelimb-related motor cortex neurons in behaving primates. Journal of Neurophysiology.
- [47] Smith, W. S. and Fetz, E. E. (2009b). Synaptic Linkages Between Corticomotoneuronal Cells Affecting Forelimb Muscles in Behaving Primates. Journal of Neurophysiology, 102(2):1040–1048.
- [48] Song, S. and Abbott, L. (2001). Cortical development and remapping through spike timing-dependent plasticity. Neuron.
- [49] Song, S., Miller, K., and Abbott, L. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature Neuroscience.
- [50] Truccolo, W., Hochberg, L. R., and Donoghue, J. P. (2010). Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. Nature Publishing Group, 13(1):105–111.
- [51] Van Rossum, M. and Bi, G. Q. (2000). Stable Hebbian learning from spike timing-dependent plasticity. The Journal of ...
- [52] Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., and Gerstner, W. (2011). Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. Science (New York, NY), 334(6062):1569–1573.
- [53] Wang, H.-X., Gerkin, R. C., Nauen, D. W., and Bi, G.-q. (2005). Coactivation and timing-dependent integration of synaptic potentiation and depression. Nature Publishing Group, 8(2):187–193.